ccle - H02962 - 9 - P 002614

Vol. 40 No. 1 (New Series No. 185) January 1980

ANALYSIS

Edited by
Christopher Kirwan

9

CONTENTS

Against a defence of cluster theories	Michael McKinsey
On ε and ∃	N. Tennant
A reply to Tennant	Christopher Peacocke
Entailment and substitution instances	A. J. Dale
Reply to Dale	Michael Clark
The impossibility of liberal egalitarianism	Albert Weale
Mackie's interpretation of Hume	H. M. Robinson
Can we get rid of events?	Irving Thalberg
Actions and bodily movements: another move	David W. D. Owen
Killing and average utility	Michael Neumann
Phenomenological specimenism	J. M. Hinton
Chisholm's definition of the evident	Wayne W[illegible]an
Goldman on epistemic conjunction	Fred Chernoff
Cartesian madness	Michael Welbourne
For want of a nail	E. J. Lowe
The transitivity of counterfactuals and causation	J. L. Mackie
Aristotle on the sea-battle: a clarification	Malcolm F. Lowe
The self from time to time	D. H. Mellor
Proper names and definite descriptions with widest possible scope	James Hudson *and* Michael Tye

CALCUTTA UNIVERSITY CENTRAL LIBRARY

BASIL BLACKWELL . ALFRED STREET . OXFORD

Ludwig Wittgenstein

Wittgenstein's Lectures 1930–1931

Edited by Sir Desmond Lee and John King

The lectures and discussions recorded in this volume are in the main based on notes taken by Sir Desmond Lee and John King in the period from January 1930 to May 1932. They correspond roughly with the period when Wittgenstein was producing *Philosophical Remarks, Philosophical Grammar,* and *The Blue and Brown Books,* and deal with such topics as the nature of a proposition, the relation of proposition to fact, the rules of grammar, infinity, the nature of proof, negation, symbols and thought.

120 pages, about £8

PRC1.4

Wittgenstein's Lectures, Cambridge 1932–1935

Edited by Alice Ambrose, from the notes of Alice Ambrose and Margaret MacDonald

The lectures which Wittgenstein gave in Cambridge between 1932 and 1935 represented a radical departure from the work he had published in the *Tractatus*.

Topics which Wittgenstein took up during this period include: the meaning of a word, its relation to rules for the word's use, general ideas, understanding as a disposition and as a conscious state, propositions, negation, rules of grammar and their relation to fact, and language games.

256 pages, about £10

Wittgenstein: Notebooks 1914–1916

Second Edition

Edited by G. E. M. Anscombe

In this new and substantially revised second edition a new appendix containing photographs of Wittgenstein's original book are included. A number of corrections have been made throughout the book and some new notes added. In addition to a new preface by Elizabeth Anscombe, the book contains a substantial index by Professor E. D. Klemke.

240 pages, £8.95

ANALYSIS 40.1 JANUARY 1980

105
An 13

AGAINST A DEFENCE OF CLUSTER THEORIES

By MICHAEL MCKINSEY

IN [2], Donald Nute has attempted to defend a version of the cluster theory of names against criticisms made by Saul Kripke in [1]. One of the tasks Nute sets for himself is that of defending the following thesis against Kripke's objections:

> (6) The statement, 'If *X* exists, then *X* has most of the ϕs' expresses a necessary truth (in the idiolect of the speaker),

where the ϕs are the properties which the speaker in question believes are satisfied by the referent of the name '*X*' in question. The sentence (6) is Kripke's way of writing a principle which is apparently endorsed by John Searle in [3]. In this note I wish to show that Nute's defence of (6) is unsuccessful.

Nute correctly points out that what we should be interested in constructing is a theory of reference for particular uses, or tokens, of proper names. So he makes the reasonable proposal that we reformulate (6) as

> (6′) Any token of the sentence 'If *X* exists, *X* has most of the ϕs' which *A* could author using the two tokens of '*X*' in the same way as [*X*] was used would express a necessary truth (in the idiolect of *A*),

where [*X*] is understood to be any token of a proper name '*X*' uttered by a speaker *A* as a proper name, and the ϕs are those properties which *A* believes a putative bearer of [*X*] would have.

Kripke gives an argument against (6) (see [1], p. 279), and Nute attempts to show that this argument is misdirected as against either (6) or (6′). It is crucial to Kripke's argument that we interpret sentences of the form

> (1) If *X* exists, then *X* has most of the ϕs

('*X*' a proper name) as being necessarily true just in case: for every possible world w and individual x, if x is the referent of '*X*' in the actual world and x exists in w, then x has most of the ϕs in w. Then Kripke's point is that instances of (1) are not necessarily true on this interpretation, since the actual referent of a name like 'Aristotle' might exist without possessing most of the properties which are associated with the name.

But Nute objects that to interpret sentences of the form (1) this way is to misinterpret (6) and (6′):

> . . . but that is not what either (6) or (6′) says. Instead (6′), for example, says it is necessarily true that in every world in which *A*'s statement that *X* exists would be true, *A*'s statement that *X* has most of the ϕ's would also be true. But the conditions for *A*'s statement that *X* exists being true are not that the thing *A* actually refers to by [*X*], if anything, exists; rather, *A*'s statement that *X* exists is true in exactly those worlds in which the conditions for *A*'s having succeeded in referring by means of [*X*], as outlined in Theses (2′) and (3′), are satisfied. We would have to determine the truth of (6′) in a particular example by looking at those worlds in which *A* actually refers by means of [*X*] and seeing whether his referent in those worlds has a weighted most of the ϕ's. But if (2′) and (3′) are correct, it would be impossible for *A* to refer by means of [*X*] in any possible world to something which didn't have a weighted most of the ϕ's. The sort of claim Kripke has attributed to sense theories may follow from certain remarks of Searle or some other theorist, but they do not follow from this interpretation of either (6) or (6′). (pp. 47–48.)

But which interpretation of (6′) is it that Nute is claiming can escape Kripke's objection? His words strongly suggest that he has in mind the principle

> (2) For any possible world w: if in w *A* utters a token α of '*X*' in a token *T* of '*X* exists' and in w *A* uses α in the same way as he uses [*X*] in w_0, then *T* is true in w only if the referent of α in w has most of the ϕs in w,

where w_0 is the base world in which *A* utters [*X*] and associates the ϕs with [*X*]. (Here we must assume, as Nute does, that *A* uses α in w "in the same way" as he uses [*X*] in w_0 only if the conditions for successful reference of α in w are the same as the conditions for successful reference of [*X*] in w_0. On the view which Nute is defending, this implies that *A* associates the ϕs with α in w and that an object x is the referent of α in w only if x satisfies most of the ϕs in w.)

Now it is true that Kripke's argument against (6) has no effect on (2). That is simply because (2) says nothing about any statements which *A* would make being necessarily true. In particular, it does not follow from (2) that

> (3) For any possible world w: if in w *A* utters tokens α and β of '*X*' in a token *T* of 'If *X* exists, then *X* has most of the ϕs' and in w *A* uses both α and β in the same way as he uses [*X*] in w_0, then *T* expresses a necessary truth in w.

But (3) says (in possible-world jargon) exactly what (6′) says. Hence (2) does not entail (6′) and so (2) is not at all the same principle as (6′). Thus *if* Nute is merely pointing out that Kripke's argument against (6) has no effect on (2), he has not said anything relevant to the defence of (6′).

But perhaps Nute is making a different point. His chief objection to Kripke's argument seems to be that Kripke's appeal to his interpretation of (1) is somehow *out of place* or, perhaps, *unfair* in an argument against (6). But it is hard to see why Nute thinks that this is so, if he does. The suggestion that Kripke, by interpreting (1) in the way he does, has somehow misinterpreted (6) is not to the point, it seems to me. If Kripke's interpretation of sentences having the form of (1) is correct (that is, if this is the interpretation or meaning which speakers' utterances of these sentences actually have), then (6) is just false. So if Kripke's interpretation of (1) is allowable in an argument against (6), then Kripke has given an allowable (fair) argument against (6). The only question then is whether or not Kripke's use of his interpretation of (1) is thus allowable.

One point which Nute may be making and which might seem relevant to this question is that a proponent of (6) or (6′) could be expected to maintain that speakers' utterances of instances of (1) do in fact have a *different* interpretation from the one Kripke suggests. This different interpretation is that where the ϕs are associated with the tokens of '*X*' in a token *T* of 'If *X* exists, then *X* has most of the ϕs', *T* is true at a possible world w just in case either no individual uniquely has most of the ϕs in w or some individual does uniquely have most of the ϕs in w. On this interpretation, of course, *T* is necessarily true. However, this fact only goes to show that proponents of (6) or (6′) might not *agree* that Kripke's interpretation of (1) is the correct one (no doubt because if it were correct, their view would be false!). But this is obviously *not* a good reason for disallowing Kripke's use of his interpretation of (1).

Since I can see no further points which Nute might be making in defence of (6′), I conclude that his defence of (6′) is unsuccessful.

But the question remains, is Kripke being fair to his opponents if he uses his interpretation of (1) in an argument against (6′)? It will help to answer this question, if we can see the role which this interpretation would play in an argument which Kripke could naturally be expected to give against (6′). Let α be a token of 'Aristotle' used in a token *T* of, say, 'Aristotle taught philosophy', where *T* is uttered by a speaker *S* in the actual world w_0. Suppose also that *S* associates the ϕs with α and that at least one conjunction of properties which consists of most of the ϕs is a complex *contingent* property, where a property *F* is contingent just in case it is possible for an object to have *F* and necessary that any object which has *F* might exist without having *F*. Finally, suppose that α *rigidly designates* a given individual *a* in w_0; that is, α designates *a* in w_0 and for every possible world w and individual *b*, α designates *b* in w if and only if *b* exists in w and $b = a$. Now is (6′) satisfied in this case? The argument that it is *not* satisfied is very simple.

Suppose now that in a world w S utters tokens β and γ of 'Aristotle' in a token T^* of 'If Aristotle exists, then Aristotle has most of the ϕs', where S uses both β and γ in w *in the same way* as he used α in w_0. Also let these tokens designate ***a*** in w, just as α does in w_0. Now according to (6′) T^* expresses a necessary truth in w. But this result is false, since using β and γ *in the same way* as α was used in w_0 requires that these tokens of Aristotle be rigid designators *too*, and so T^* will be false in any possible world w* in which *a* exists and yet fails to have most of the ϕs (there is at least one such world, since at least one conjunction consisting of most of the ϕs is contingent, by hypothesis). Hence, (6′) is false.

The role of the Kripkean interpretation of sentences of the form (1) in the above argument is clear. The use of this interpretation, if it is justified, is justified completely by the initial assumption that the token α of 'Aristotle' that occurs in the token T of 'Aristotle taught philosophy' is a rigid designator, and this is an assumption that Kripke would of course make. It is worth noting that the argument against (6′) does *not* assume that any speaker of any actual language *ever* actually gives a Kripkean interpretation to any sentence of the form (1). Thus Kripke's use of this interpretation is only as controversial as his view that names are very often used as rigid designators. (Strictly, (6′) is shown false just given the weak assumption that at least one token of a name rigidly designates an object. But for the argument to be really effective against (6′) we need the assumption that names are very often used as rigid designators; for otherwise we could not prevent the proponent of (6′) from just slightly weakening his view and holding that, at any rate, names are *typically* used so as to satisfy (6′).)

So the question of whether Kripke has a right to use his interpretation of (1) in an argument against (6′) reduces to the question of whether Kripke has a right to use his view that names are very often used as rigid designators. I think it is obvious that Kripke does have this right. He has said enough in [1] to make his view intuitively plausible, and so he can justifiably use it in arguments against other views. Of course, this is not to say that Kripke's view is true or that the evidence Kripke has given enables us to know that it is true. But it is to say that Kripke is not being unfair to his opponents, if he uses his view to argue against a thesis like (6′). Moreover, since Kripke's view about rigid designators does yield an allowable and valid argument against (6′), a successful defence of (6′) *must* consist of an argument against Kripke's view, or at least of an argument to the effect that the view is unsupported by the evidence. Nute, of course, has given no such argument and consequently he has not successfully defended (6′).

Wayne State University © Michael McKinsey 1980

REFERENCES

[1] Kripke, Saul A., 'Naming and Necessity,' in *Semantics of Natural Language,* edd. D. Davidson and G. Harman, Dordrecht, D. Reidel, 1972.
[2] Nute, Donald E., 'Proper Names: How to Become a Causal Theorist While Remaining a Sense Theorist,' *Philosophia,* VIII, 1 (1978), pp. 43–57.
[3] Searle, John R., 'Proper Names,' *Mind,* LXVII (1958), pp. 166–73.

ON ε AND ∃

By N. TENNANT

IN his paper 'A reply to Bjürlof's objection' in ANALYSIS 38.3, June 1978 pp. 122–4, Peacocke considers extending his 'criterion of logical constanthood that applies . . . to primitive expressions that are assigned semantic properties directly by the axioms of a truth theory' to apply to non-primitive expressions that are definable in terms of the primitive ones. He summarizes his discussion as follows:

> '∃' defined in terms of 'ε' may be a logical constant relative to a truth theory for an object language with an ontology of natural numbers; but equally my criterion pronounces it as such. On a narrower notion of logical constanthood, one might require that an expression be a logical constant in the original sense with respect to *all* truth theories that differ from one another only in the richness of the ontology of the object language. In that sense, Hilbert's ε-operator is not plausibly a logical constant, nor do I make it one, nor is '∃' so generally definable in terms of 'ε'.

Note that Peacocke's criterion cannot be applied to yield an affirmative or negative answer if the expression concerned is primitive but is not 'assigned semantic properties directly by the axioms of a truth theory'. This particularly affects the ε-operator, since Peacocke claims dogmatically that the natural axiom for ε in a homophonic theory, namely

$$s(\varepsilon x A)=\varepsilon\alpha(s(x/\alpha) \text{ satisfies } A)$$

has its truth conditions 'left unfixed' in the metalanguage—a metalanguage, remember, in which ε is primitive.

It is Peacocke's final claim, however, that is most seriously in error, given a proper understanding of what it is for one operator to be 'generally definable' in terms of another (in this connection, see Smiley, 'The independence of connectives', *Journal of Symbolic Logic* 27, 1962). I wish to claim

(i) Hilbert's ε-operator *is* plausibly a logical constant. Either the natural axiom above should make it so, or Peacocke's criterion must be admitted to have too narrow a range of application, and

(ii) ∃ *is* generally definable in terms of ε.

Take a first order language without quantifiers. Let ε be the only variable-binding operator in the vocabulary, forming terms from formulae. Adjoin to the usual rules of natural deduction for the connectives and identity (as given, for example, in my *Natural Logic*, Edinburgh, 1978) the single rule

$$\frac{A^x_t}{A^x_{\varepsilon xA}}$$

This rule does not extend the deducibility relation among ε-free sentences.

The existential quantifier can then be defined as follows:

$$\exists xA =_{df} A^x_{\varepsilon xA},$$

so the rule above is the introduction rule for ∃. The corresponding elimination rule:

$$\cfrac{\exists xA \quad \begin{matrix}\overline{A^x_a}\\ \vdots \\ B\end{matrix}}{B}$$

(where a does not occur in $\exists xA$, B or any assumption other than A^x_a on which the upper occurrence of B depends) comes out in the wash simply as

$$\begin{array}{l} A^x_{\varepsilon xA} \\ \vdots\ a/\varepsilon xA \\ B \end{array}$$

where the slash notation indicates that throughout the proof free occurrences of a are replaced by occurrences of εxA, re-labelling bound occurrences of names as necessary. (For the notion of free and bound occurrences of names in proofs, see *Natural Logic*, p. 66.)

The universal quantifier can also be defined:

$$\forall xA =_{df} A^x_{\varepsilon x-A}$$

The introduction rule for ∀:

$$\frac{\begin{matrix}\vdots \\ A\end{matrix}}{\forall xA^a_x}$$

(where a does not occur in any assumption on which A depends) is derived thus:

$$\begin{array}{l} \vdots\ a/\varepsilon x-A \\ A^a_{\varepsilon x-A} \end{array}$$

and the corresponding elimination rule:

$$\frac{\forall x A}{A^x_t}$$

is derived by the proof

$$\dfrac{\dfrac{A^x_{\varepsilon x - A} \qquad \dfrac{\dfrac{\overline{}\,(1)}{-A^x_t}}{-A^x_{\varepsilon x - A}}}{\Lambda}}{A^x_t}\,(1)$$

With appropriate adjustments, these considerations hold for a universally free system as well.

On the strength of these observations alone ε must be regarded as a logical constant. It is governed by a single, simple rule and is the sole operator appearing in quite legitimate definitions of the ordinary quantifiers. Any criterion of logical constanthood must square with the intuition that a defining expression is a logical constant if it can feature as the only primitive expression in definitions of other logical constants. Aversion to the natural axiom for ε in a homophonic truth theory must be due to lack of familiarity with an ε-language.

Unless ε is a logical constant, it is nothing. Peacocke must then explain how we get something out of nothing.

University of Edinburgh

© N. TENNANT 1980

A REPLY TO TENNANT

By CHRISTOPHER PEACOCKE

WHAT Neil Tennant argues to be a logical constant ('On ε and ∃', ANALYSIS 40.1, January 1980) is not what I deny to be one. Suppose that in a natural deduction system we give this introduction rule for Hilbert's ε-operator (I follow Tennant's rule and notation):

$$\frac{A^x_t}{A^x_{\varepsilon x A}}\ .$$

This, given the background assumption that it is the strongest rule the introduced constant satisfies, is meant to be a complete specification of the meaning of 'ε'. But the formation rules of the extended language including 'ε' mean that there are more degrees of freedom for possible occurrences of 'ε' than are settled by its introduction rule: there are many sentences of the form $B(\varepsilon x A)$ which do not have determinate truth conditions though the truth conditions of all sentences in the ε-free fragment are determinate.

If we changed our notation, this indeterminacy would disappear. Our introduction rule essentially introduces a unary operator on the predicate $A(x)$. Given that natural deduction is to preserve truth in a familiar way, we can generalize arguments suggested by Ian Hacking and myself to show that such an operator must have precisely the satisfaction conditions of the existential quantifier (see section XVI of Hacking's 'What is Logic?', *Journal of Philosophy* 76 (1979) 285–319, and the reference therein; and my 'What is a Logical Constant?', *Journal of Philosophy* 73 (1976), at p. 232). But the classical existential quantifier is a logical constant by my original criterion. For these reasons, I am in complete agreement with Tennant that 'ε' so introduced should be regarded as a logical constant in those sentences whose truth conditions are fixed, relative to those of the ε-free fragment, by the above rule. What I was denying in the reply to Bjurlöf is that there is an acceptable general method of fixing up the cases which the rule does not settle, and which still makes 'ε' a logical constant.

There remains a disagreement over the natural homophonic axiom:

$$s^*(\ulcorner \varepsilon v A \urcorner) = \varepsilon x(s(v/x) \text{ satisfies } A).$$

Tennant implies that it is acceptable if 'ε' is primitive in the metalanguage, and writes 'Aversion to ... [it] must be due to lack of familiarity with an ε-language' (p. 7). But here there is a dilemma. Either 'ε' has the same sense in the metalanguage as in the object language, or it does not. If it does not, the homophonic axiom is irrelevant: for even if it does have determinate truth conditions, it is not an acceptable specification of meaning. If on the other hand 'ε' does have the same sense in the two languages, that sense is fully specified in the above introduction rule, now stated for metalanguage sentences. This rule certainly does not determine which object, for given s, v and A, is to be $\varepsilon x(s(v/x)$ satisfies $A)$. Many different assignments are compatible with the rule.

Perhaps someone sympathetic to Tennant's attitude might say this: 'Forget about assigning objects to ε-terms. Just accept the axiom, and take it that predicate position inside the ε-operator is extensional; you will then find that with the other axioms of a homophonic truth theory you will be able to prove disquotational T-sentences for object language sentences containing "ε".' The second sentence of this comment is technically correct; and the first might be elaborated by taking $B(\varepsilon x A)$ as an unusual way of writing a binary predicate operator $\mathrm{E}x(A,B)$. But it is not sufficient defence of the homophonic axiom. Suppose the binary operator 'E' has its meaning fully specified by the single rule

$$\frac{A^x_t}{\mathrm{E}x(A,A)}$$

The homophonic axiom is then written

$$s \text{ satisfies } \ulcorner \mathrm{E}v(A,B)\urcorner \equiv \mathrm{E}x(s(v/x) \text{ satisfies } A,\ s(v/x) \text{ satisfies } B).$$

But *this* notational change does not produce determinacy of truth conditions; again this axiom is not a sentence whose truth conditions are settled by the rule for 'E' together with the senses of its other components. All we can tell from the rule is that (in the case of a closed sentence) $\ulcorner \mathrm{E}v(A,A)\urcorner$ is true iff something satisfies A. Whatever the strengths of the doctrine that axioms in a truth theory are justified by their consequences at the level of T-sentences, it cannot be so strong as to support the view that we can employ axioms which do not say anything.

New College, Oxford

© CHRISTOPHER PEACOCKE 1980

ENTAILMENT AND SUBSTITUTION INSTANCES

By A. J. Dale

IN ANALYSIS 39.1 (January 1979) Michael Clark proposes a new variant of Geach's substitution instance account of entailment in order to excise certain entailments which he claims involve irrelevant components. His definition is as follows (bottom of p. 33):

> $A \rightarrow B$ iff $A \supset B$ is a substitution instance of some tautology with no non-contingent well formed proper parts.

Now this new definition is open to all the objections that I gave to Geach's original definition and which Geach has subsequently modified (not successfully, as I have shown).[1] Since Clark mentions that there are possible objections it is perhaps unfair to tax him with these same arguments and examples. But there are other examples which cannot be ignored in the light of his further generalization of the definition and the use he makes of it.

According to his definition the following entailments do not hold, if each logical connective is taken as primitive:

$$(p \equiv {\sim}p) \rightarrow (p \,.\, {\sim}p)$$
$$(p \supset p) \rightarrow (p \equiv p)$$

and the following two do hold:

$$(p \equiv {\sim}p) \rightarrow ((p \,.\, {\sim}p) \vee (p \,.\, {\sim}p))$$
$$((p \supset p) \,.\, (p \supset p)) \rightarrow (p \equiv p).$$

That is, we have examples where $A \rightarrow B \vee B$ but not $A \rightarrow B$ and where $A \,.\, A \rightarrow B$ but not $A \rightarrow B$.

These new examples, in addition to casting doubt on the acceptability of his definition, prove pertinent for his further remarks. In order to make sense of the claim that in Quine's system in *Mathematical Logic* 'the Burali-Forti paradox was found to arise, although Cantor's paradox was apparently avoided' Clark suggests an appropriate generalisation of his definition to quantificational logic (p. 34 footnote 1):

> $A \rightarrow B$ iff the closure of $A \supset B$ is a substitution instance of a valid closure of some wff a closure of no proper part of which is valid or inconsistent (in the sense of logically false).

Now, Clark claims this defined entailment relation should hold between a conjunction of Quine's axioms and the contradiction involved in the Burali-Forti paradox, but not between any such conjunction and the contradiction in Cantor's paradox. Whether this situation does indeed obtain remains to be seen but by using variants of the above examples it can be shown that the programme is misdirected.

---

[1] P. T. Geach, 'Entailment' *Philosophical Review,* 1970 pp. 237–239. A. J. Dale, 'Geach on Entailment', *Philosophical Review,* 1973 pp. 215–219. P. T. Geach, 'On Entailment', ANALYSIS 35.6, June 1975 pp. 186–187. A. J. Dale, 'Geachian Entailment', *Philosophical Review* 1978 pp. 423–426.

As a counterexample let me take a simpler example. It is undisputed that in a set-theory based upon the naive axiom schema of comprehension,

$$(\mathrm{E}y)(x)(x \varepsilon y \equiv Fx)$$

Russell's paradox arises. If Clark's theory is correct we should expect that some conjunction of instances of this schema must entail (under his definition) Russell's paradox. What then is Russell's paradox? Here there is a problem. There are many ways in which we might wish to express this paradox but one which uncontrovertably does express it is $(\mathrm{E}x)(x \varepsilon x \,.\, \sim x \varepsilon x)$. Normally the derivation of this wff from the axiom schema uses only one instance of the schema, *videlicet*

$$(\mathrm{E}y)(x)(x \varepsilon y \equiv \sim x \varepsilon x)$$

but it is not the case that

$$(\mathrm{E}y)(x)(x \varepsilon y \equiv \sim x \varepsilon x) \supset (\mathrm{E}x)(x \varepsilon x \,.\, \sim x \varepsilon x)$$

is a substitution instance of any wff with no non-contingent proper parts. So are we to say that Russell's paradox does not arise if we adopt the naive axiom of comprehension? Clearly there is something wrong with a definition which gives this result. We can further confound matters if we look at what might be considered an alternative way of stating Russell's paradox: $(\mathrm{E}x)(x \varepsilon x \equiv \sim x \varepsilon x)$. We might expect that this wff entails $(\mathrm{E}x)(x \varepsilon x \,.\, \sim x \varepsilon x)$ but on Clark's definition it does not. However, our axiom of comprehension does entail $(\mathrm{E}x)(x \varepsilon x \equiv \sim x \varepsilon x)$ since

$$(\mathrm{E}y)(x)(x \varepsilon y \equiv \sim x \varepsilon x) \supset (\mathrm{E}x)(x \varepsilon x \equiv \sim x \varepsilon x)$$

is a substitution instance of

$$(\mathrm{E}y)(x)(Fxy \equiv Gx) \supset (\mathrm{E}x)(Fxx \equiv Gx)$$

which has no non-contingent proper parts. It matters then what we consider to be '*the* contradiction involved in the . . . paradox'. Normally we are quite happy to move from $(\mathrm{E}x)(x \varepsilon x \equiv \sim x \varepsilon x)$ to $(\mathrm{E}x)(x \varepsilon x \,.\, \sim x \varepsilon x)$ and so perhaps not worry as to exactly when we have reached the contradiction, but, if we were to take Clark's definition, the explicit contradiction is not entailed by the axiom of comprehension.

Of course, this example is dependent on the earlier examples taken from the propositional calculus since basically it reduces to the fact that

$$(p \equiv \sim p) \rightarrow (p \,.\, \sim p)$$

is false on Clark's definition. Until such difficulties are cleared up it is unwise to generalize to the predicate calculus. The use of the notion of substitution instances in solving the problems of entailment has proved trickier than its propounders thought, hence Geach's modification. Just as Geach's modification was found not to be an improvement after all, so Clark's modification and generalisation represent neither an improvement nor an advance. Perhaps after a good run of twenty years the notion of substitution instance should be given a timely retirement.

University of Hull © A. J. DALE, 1980

REPLY TO DALE

By MICHAEL CLARK

DALE is wrong when he claims (ANALYSIS 40.1, January 1980) that the explicit contradiction of Russell's paradox, $(\exists x)(x \in x \,.\, {\sim} x \in x)$, is not entailed in my suggested sense (ANALYSIS 39.1, January 1979) by any instance of the naive axiom schema of comprehension, $(\exists y)(x)(x \in y \equiv Fx)$ (I follow Dale here in treating both $\equiv$ and . as primitive). Granted, as he shows, it is not thus entailed by $(\exists y)(x)(x \in y \equiv {\sim} x \in x)$; but it does not follow that it is not entailed in that sense by any instance whatsoever of the schema. It is entailed, for example, by $(\exists y)(x)(x \in y \equiv (x \in x \supset {\sim} x \in x))$, obtained by substituting the stencil $(1) \in (1) \supset {\sim}(1) \in (1)$ for F in the schema, since

$$(\exists y)(x)(x \in y \equiv (x \in x \supset {\sim} x \in x)) \supset (\exists x)(x \in x \,.\, {\sim} x \in x)$$

is a substitution instance, obtained by substituting $(1) \in (2)$ for F and ${\sim}(1) \in (I)$ for G, of the valid wff

$$(\exists y)(x)(Fxy \equiv (Fxx \supset Gx)) \supset (\exists x)(Fxx \,.\, Gx),$$

the closure of every wf proper part of which is contingent.

University of New Orleans and
University of Nottingham

© MICHAEL CLARK 1980

THE IMPOSSIBILITY OF LIBERAL EGALITARIANISM

By Albert Weale

I. The Problem

BY 'liberal egalitarianism' I mean a social theory which states that a more equal distribution of benefits is ethically preferable to a less equal distribution, unless there is a liberal reason to the contrary. One such reason is that a certain policy leaves everyone with a larger total stock of benefit than they would possess under equality. Following the usage of economics, I shall say that such distributions are Pareto-superior to their more equal alternatives. A number of writers have interpreted egalitarianism to mean that there is a *prima facie* burden of proof on the proposer of inequality, and Rawls (1972, p. 62) has argued that at the basis of his theory there is the presumption in favour of equality unless there is a Pareto-superior alternative. Raz (1978, p. 338) identifies this version of the equality principle with a strict form of egalitarianism. I shall argue that this social theory is based on a deep inconsistency. The 'liberal' modification to egalitarianism introduced by the use of the Pareto principle is more far-reaching than is normally thought to be the case.

Suppose a formulation of the equality principle, E, which is capable of ranking the elements in a set of alternative, feasible distributions $\{x_1, x_2, \ldots, x_n\}$. If one alternative is preferred to another by this criterion, write '$x_1 >_E x_2$'. The ordering this criterion yields is both transitive and complete: if $x_1 >_E x_2$ and $x_2 >_E x_3$, then $x_1 >_E x_3$; and for any pair of alternatives the equality rule will produce a ranking either of preference (e.g. $x_1 > x_2$) or indifference (e.g. $x_1 = x_2$). It is implicit in this formulation that the egalitarian's preference is applicable both to distributions of a given total of benefit and to distributions where the total benefit varies. This is so as to enable the egalitarian to judge whether an increase in the average level of benefit in society is consistent with his opinion of an ethically preferred distribution. In practice, for example, the egalitarian will wish to see whether the benefits of economic growth outweigh any resulting increase in relative deprivation. In this context, the simplest assumption to adopt is that the egalitarian will measure inequality in such a way as to factor out any variations in the absolute size of benefits. (Thus one common measure of income inequality, the Gini coefficient, is standardised for different average levels of income, cf. Sen, 1973, p. 31). An egalitarian therefore will be strictly indifferent between two alternatives with the same degree of inequality but different total levels of benefit. To choose between them, he will have to supplement the principle of equality by some rule of maximization. The Pareto principle is the simplest and least ethically controversial of such rules.

The Pareto principle yields a P-ranking of alternatives according to the test of whether everyone is better off in one alternative than in another. So '$x_1 >_P x_2$' means that everyone is better off in x_1 than in x_2. Since not all feasible public policies result in everyone being better off, the Pareto principle will not normally yield a complete ordering. However, a P-ranking will be transitive: '$x_1 >_P x_2$' and '$x_2 >_P x_3$' will imply '$x_1 >_P x_3$'.

Sometimes the equality and Pareto principles conflict: according to one criterion $x_1 >_E x_2$ and according to the other $x_2 >_P x_1$. For example, economic growth may increase everyone's level of benefit, but also increase the gap between the worst off and best off. By virtue of his social theory, the liberal egalitarian will prefer the P-ranking to the E-ranking in such cases. So '$x_1 >_E x_2$' and '$x_2 >_P x_1$' imply a final ordering '$x_2 > x_1$'. Does this liberal priority rule for reconciling two sometimes contradictory rankings produce a consistent overall liberal egalitarian evaluation? The answer is, in general, no. It can be shown easily why not. Suppose an E-ranking, $x_1 >_E x_2 >_E x_3$, and a P-ranking, $x_3 >_P x_1$. Then by the liberal priority rule, $x_3 > x_1$, and by the undominated egalitarian principle, $x_1 > x_2$. But these two pair-wise comparisons do not imply $x_3 > x_2$, which is what would be needed if liberal egalitarianism were to be consistent. There is no non-arbitrary way, within liberal egalitarian principles, for selecting which of three alternatives is best, since the two independent sets of transitive rankings do not combine into an overall transitive ordering.

One obvious solution to this problem is to strengthen the liberal priority rule to make an ordering emerge. There are two ways in which this might be done. (i) The liberal component of the theory might be strengthened by requiring that any alternative x_1 which is Pareto-superior to x_2 be ethically preferred to x_3, if x_3 is inferior to x_2 on *either* the P-ranking *or* the E-ranking. (ii) A second approach is to strengthen the egalitarian aspect of the theory by requiring that any alternative which is egalitarianly superior to and not Pareto inferior to a second alternative which Pareto dominates a third alternative is also ethically preferred to that third alternative. Both approaches build in the necessary element of transitivity into the liberal egalitarian theory. To illustrate their application consider the following E-rankings and P-rankings to be combined into a liberal egalitarian ordering: $x_1 >_E x_2, x_2 >_E x_3, x_3 >_E x_4, \ldots, x_{n-1} >_E x_n$ and $x_3 >_P x_1, x_1 >_P x_4$. Approach (i) would yield the following overall ordering: $x_3 > x_1,\ x_1 > x_2,\ x_2 > x_4, \ldots, x_{n-1} > x_n$. Approach (ii) on the other hand gives: $x_2 > x_3,\ x_3 > x_1,\ x_1 > x_4, \ldots, x_{n-1} > x_n$.

Both approaches involve difficulties however. In particular they both violate what may be termed, following statistical decision theory, 'the independence of irrelevant alternatives'. In decision theory this condition is interpreted to require that, if an act is non-optimal for a decision problem, it cannot be made optimal by adding alternative prospects over which a choice must be made (Luce and Raiffa, 1957, p. 288; cf. Chernoff, 1954, p. 426). By extension one can argue that if an alternative is not regarded as ethically preferable to another alternative, it cannot be made such by the addition of a third, independent, and hence irrelevant, alternative. Yet this requirement is necessarily violated by both strengthened versions of liberal egalitarianism. In each case the ranking of two alternatives (in the example x_1 and x_2) may depend on how they are each ranked relative to some third alternative (in the example, x_3).

How important is it to respect the independence of irrelevant alternatives? By violating this requirement we may invert an ordering of alternatives not because anything has changed in respect of those alternatives, but because we had introduced or deleted a third state in the scope of choice. So an element x_2 might be the best element in a given triple $\{x_1, x_2, x_3\}$ but not the best element in a pair-wise comparison from this triple, e.g. $\{x_1, x_2\}$. The rule that, if any element is the best element in a set, then it should also be the best element in any proper subset of that set, really is an attractive rationality requirement as Sen (1970, p. 81) pointed out, and one would not want to violate it lightly.

Moreover, to violate the independence of irrelevant alternatives is to base one's system of social choice not on binary comparisons, but on over-all comparisons which may issue in binary comparisons. However, using an argument of Murakami (1968, p. 107), it can be shown that this approach reverses the true logical relationship between binary and higher-order comparisons. To see why, suppose a choice between four alternatives $\{x_1, x_2, x_3, x_4\}$ and a choice rule based on ternary comparisons. Take the triple $\{x_1, x_2, x_3\}$. The choice rule will define an ordering between each pair in the triple, which can be expressed as the following choice functions: $C'(x_1, x_2)$, $C'(x_2, x_3)$, and $C'(x_1, x_3)$. Similarly, the rule will define another set of choice functions for the triple $\{x_2, x_3, x_4\}$ as follows: $C''(x_2, x_3)$, $C''(x_3, x_4)$ and $C''(x_2, x_4)$. However, $C'(x_2, x_3)$ and $C''(x_2, x_3)$, the product of two distinct ternary comparisons, can only be consistent provided that the ternary comparisons are reducible to pair-wise comparisons. For if $C'(x_2, x_3)$ and $C''(x_2, x_3)$ are identical, the choice concerning the pair $\{x_2, x_3\}$ is independent of the orderings assigned to the elements for other pairs in the choice set. Hence, to arrive at a consistent decision concerning any pair, we need an independent pair-wise comparison of the elements in that pair. The same argument can be generalised to quaternary and higher-order comparisons. So any rehabilitation of liberal egalitarianism which involves choosing between a pair of alternatives by looking at their relation to a third will encounter the difficulty that no consistent binary choice system could be expected to emerge from such a procedure.

II. Explanation

The difficulty of formulating priority rules adequately to adjudicate between the equality and Pareto principles points to a deep inconsistency in the theory of liberal egalitarianism. How can it be accounted for? Its basic source, I suggest, is that equality is a 'strongly organic' principle, a special case of what Nozick (1974, p. 209) calls 'organic' patterned principles. Organic principles fail to satisfy the deletion condition, which requires that deleting (in imagination) some people and their shares from an ethically preferred distribution leaves the remaining distribution ethically preferred. Lexical maximin, average utility and equality are all principles failing to satisfy this condition. Under each of these principles, deleting one person from a distribution might mean that another distribution, previously regarded as inferior, was now regarded as ethically superior.

Unlike the equality rule, however, the organic principles of maximising average utility or lexically maximising the well-being of the least well-off positions are consistent with the Pareto principle. There is, for example, no conflict between a P-ranking and maximising average utility, since any overall improvement in the level of benefits will lead to a higher average level of benefit. Similarly, the rankings produced by lexical maximin are a sub-class of the class of P-rankings for a set of distributions. These two principles, then, may be termed 'weakly organic'. (Note, incidentally, the distinction between lexical maximin, which allows inequalities if they fail to harm the least well-off, and Rawls's (1972, p. 302) difference principle, which requires inequalities to benefit the least well-off. The latter more egalitarian principle is not Paretian in the same way as the former is).

As a strongly organic principle, equality is not Paretian at all. Individual welfare improvements are not admissible unless they equalize the overall structure of all individuals' benefit. Within this framework, it is no justification of an increment to social welfare merely to say that it improves the position of individuals relative to what it was before. Even if the position of the least well-off is improved, this might be at the cost of an increase in inequality, which would conflict with egalitarian requirements. So improving the position of the least well-off has to be conditional upon there being an improvement also in the equality measure. The equality rule is then a genuinely collective principle, in as much as the estimation of individual holdings can only be made relative to the general structure of the distribution of benefits. The lot of each individual is tied to the lot of every other individual. It is this social holism which creates the logical inconsistency with the individualistic Pareto principle.

III. Implications

1. One implication which might be drawn from the above is that the Pareto principle, even in its weak form, should be dropped as a condition of social choice. Moreover, the association of the Pareto principle with liberalism might also be denied. Sen (1970, p. 83) showed that the Pareto principle, when combined with the 'liberal' requirement that individuals be allowed to determine exclusively some aspects of their own lives, led to an inconsistency; and Sugden (1978) has argued that this inconsistency points to the illiberal nature of orthodox Paretian social choice theory. Rowley and Peacock (1975, pp. 103–33) too deny that liberals are committed to the Pareto principle.

In reply one should first note that the Pareto principle need not comprise the whole of the liberal content of liberal egalitarianism. For example the liberal egalitarian would presumably wish to impose constraints on the way in which egalitarian distributions are brought about. However, it would seem difficult for him to deny that his ordering of such constrained alternative distributions should be consistent with the Pareto principle. In the first place liberalism is usually understood as a want-regarding theory (Barry, 1965, p. 66) and this may be taken to imply at a minimum that a social decision-maker or public official should not impose his preference on a society when everyone in that society would hold themselves to be better off in some alternative. Secondly, the denial of an association between liberalism and Paretianism rests crucially on the claim that the idea of a social *ordering* is essentially illiberal (cf. Sugden, 1978). But this manoeuvre is not open to the liberal egalitarian, who wishes to preserve the idea of an ordering as the framework within which he can make judgements across alternatives. Thirdly, the weak Pareto principle is so mild that it is difficult to see how any ethical system which claimed to be individualistic could jettison it.

2. An alternative implication would be to deny that liberal egalitarianism should issue in a transitive ordering, particularly as transitivity is sometimes regarded as a strong imposition on methods for choosing social states.

However, this argument is less strong that it may seem. As originally formulated liberal egalitarianism will not even produce an ordering of social states which is acyclical, that is, where '$x_1 >_E x_2 >_E x_3$' and '$x_3 >_P x_1$' imply '$x_3 > x_1$, $x_1 > x_2$' and '$x_3 \geqslant x_2$'. The liberal priority rule might be amended, so that acyclicity were satisfied, but as with the other proposed amendments the independence of irrelevant alternatives would be violated. This suggests that the intransitivity of liberal egalitarianism is a symptom rather than a cause of the problem, particularly as acyclicity is a weaker requirement than transitivity.

Moreover, in this context transitivity is not being imposed on a social decision procedure, like a voting mechanism, but on an individual's ethical system. For voting systems transitivity may be too strong a condition, in particular if it is imposed to ensure 'the independence of the final choice from the path to it' (Arrow, 1963, p. 120) since weaker conditions than transitivity will suffice to secure this aim (Brams, 1976, p. 37). But when we are dealing with an individual's attempt to make his ethical judgments on social alternatives consistent, transitivity is an attractive property, since it amounts to little more than the requirement that an individual know his own mind all things considered.

3. A third implication, of course, is simply that equality should be dropped as a principle of social choice. Indeed the inconsistency might be used to convince somebody who held liberal values that equality was incompatible with his commitment.

The egalitarian reply to this implication may take a number of forms. It can be argued that the relevant set of social alternatives never or rarely occurs in practice, so that the inconsistency is not a practical problem. Moreover, given that other social choice devices, including majority rule for example, are flawed by certain formal deficiencies, the result might simply show that liberal egalitarianism is subject to certain difficulties, not that it faces insuperable objections. Finally, it might be argued that the two principles have different domains of application: the Pareto principle enjoining maximization of production possibilities and equality stipulating the proper distribution of consumption opportunities.

Each of these replies may be reasonable responses to the suggestion that equality should be dropped as a principle of social choice. However, the inconsistency at the heart of liberal egalitarianism ought at least to be made explicit when considering prescriptions on public policy. I for one conclude that the inconsistency suggests strong grounds against taking unreformed liberal egalitarianism as a viable social theory.[1]

University of York © ALBERT WEALE 1980

REFERENCES

Arrow, K. J. (1963) *Social Choice and Individual Values* (New Haven: Yale University Press, Second Ed.).
Barry, B. M. (1965) *Political Argument* (London: Routledge).
Brams, S. J. (1976) *Paradoxes in Politics* (New York: Free Press).

---

[1] I should like to thank Thomas Baldwin, Peter Jones, Iain McLean and Robert Sugden for their detailed comments and criticisms upon earlier versions of this paper.

Chernoff, H. (1954) 'Rational Selection of Decision Functions' *Econometrica* 22 pp. 422–43.
Luce, R. D. and Raiffa, H. (1957) *Games and Decisions* (New York: Wiley).
Murakami, Y. (1968) *Logic and Social Choice* (New York: Dover).
Nozick, R. (1974) *Anarchy, State and Utopia* (Oxford: Blackwell).
Rawls, J. (1972) *A Theory of Justice* (Oxford: Clarendon).
Raz, J. (1978) 'Principles of Equality' *Mind* 87: 347 pp. 321–42.
Rowley, C. K. and Peacock, A. T. (1975) *Welfare Economics: A Liberal Restatement* (London: Martin Robertson).
Sen, A. K. (1970) *Collective Choice and Social Welfare* (San Francisco: Holden Day Inc.).
Sen, A. K. (1973) *On Economic Inequality* (Oxford: Clarendon).
Sugden, R. S. (1978) 'Social Choice and Individual Liberty' in M. Artis and A. R. Nobay (eds.) *Contemporary Economic Analysis* (London: Croom Helm).

MACKIE'S INTERPRETATION OF HUME

By H. M. Robinson

I

MY purpose in this discussion is to challenge two points in Mackie's interpretation of Hume on necessity and causation.[1] In section one I shall dispute Mackie's reconstruction of the steps by which Hume is said to reach the conclusion that necessity does not reside in objects: in the second section I shall defend Hume against Mackie's charge that he fails to distinguish between that form of necessity which characterizes causation and that form which solves the problem of induction.

According to Mackie (p. 11), Hume's conclusion concerning causation is

(a) Causation in objects, so far as we know, is only regular succession.

This conclusion is immediately derived from two premises, (b) and (d). (I shall use Mackie's lettering for ease of reference).

(b) Necessity is in the mind, not in objects.
(d) Causal knowledge and inference, and the idea of necessary connection, arise purely from the experience of constant conjunction.

I shall argue that Hume does not make (a) depend on (b): instead I shall support the more natural interpretation that Hume first proves that causation in objects is nothing but constant conjunction and only then, in order to avoid the conclusion that we have no proper idea of necessity, tries to prove that necessity is in the mind. In order to demonstrate this I shall claim, first, that even on Mackie's reconstruction of Hume (a) can be derived without the assistance of (b): and, second, that textual evidence favours the interpretation that (a) precedes (b).

---

[1] J. L. Mackie, *The Cement of the Universe*, Oxford, 1974. D. Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, Oxford, 1964.

I agree with Mackie (p. 7 and p. 10) in attributing to Hume the following proposition:

> (e) This experience [of constant conjunction] neither reveals nor produces any necessity in objects: that is, it does not provide materials for any rational inference from cause to effect (or vice versa) in a new instance.

Propositions (d) and (e) together entail (a), for (d) with the first clause of (e) say that experience of constant conjunction gives rise to the idea of necessity, but does not reveal it to be present in objects. Thus the argument from these premises leaves it as a puzzle how necessity derives from constant conjunction, and it is this remaining puzzle which the proof of (b) is intended to remove.

Hume's succinct summary of his argument also shows this progression:

> The idea of necessity arises from some impression. *There is no impression conveyed by our senses, which can give rise to that idea.* [My italics]. It must, therefore, be derived from some internal impression, or impression of reflexion. There is no internal impression, which has any relation to the present business, but that propensity, which custom produces, to pass from an object to the idea of its usual attendant (p. 165).

In this passage Hume concludes that 'therefore' necessity must be internal only *after* concluding that it cannot have come from the external world. The same point is made slightly later (in the passage from which Mackie derives both (d) and (e) above).

> . . . We repeat to ourselves *that* the simple view of any two objects or actions, however related, can never give us any idea of power, or of a connexion betwixt them: *that* this idea arises from a repetition of their union: *that* the repetition neither discovers nor causes any thing in the objects, but has an influence only on the mind, by that customary transition it produces: *that* this customary transition is, therefore, the same with the power and necessity (p. 166).

Once the argument has reached the point that 'repetition neither discovers nor causes any thing in the object' it follows that causation in objects will contain nothing except constant conjunction; that is, (a) follows. The production of the idea by its influence in the mind is added after this conclusion.

Furthermore, Hume sometimes countenances the thesis that there is no impression of necessity, so clear is it that no such thing lies among objects.

> Thus upon the whole we may infer, that when we talk of any being, whether of a superior or inferior nature, as endow'd with a power or

> force, proportion'd to any effect; when we speak of a necessary connexion betwixt objects, and suppose, that this connexion depends upon an efficacy or energy, with which these objects are endow'd; in all these expressions, *so apply'd*, we have really no distinct meaning, and make use only of common words, without any clear and determinate ideas. But as 'tis more probable, that these expressions do here lose their true meaning by being *wrongly apply'd*, than that they never have any meaning; 'twill be proper to bestow another consideration on this subject, to see if possibly we can discover the nature and origin of those ideas, we annex to them (p. 162).

Hume then proceeds to explain how experience of constant conjunction develops in the mind a tendency to pass from one to the other. It is plain that this searching for an alternative explanation is entered upon only after the possibility that causation as it is known in objects might possess necessity has been thoroughly discredited.

II

Mackie distinguishes two types of necessity which he says Hume conflates although 'it is a sheer assumption that these two interpretations belong together' (p. 12).

> We might call *necessity*$_1$ whatever is the distinguishing feature of causal as opposed to non-causal sequences, and *necessity*$_2$ the supposed warrant for an *a priori* inference, for example, the power which if we found it in C would tell us at once that C would bring about E, or some corresponding feature which if we found it in E, would tell us at once that E had been produced by C, or some relation between one kind of thing and another which would tell us that things of those kinds could occur only in sequences, with a thing of the second kind following a thing of the first (pp. 12–13).

Mackie claims that necessity$_1$ is the key to an analysis of causation, whereas a legitimate form of necessity$_2$ would solve the problem of induction by allowing inference from past to future cases. Hume, he claims, by failing to distinguish the two forms of necessity, fails to distinguish the problem of causation from the problem of induction, and in trying to uncover causal necessity tries to make candidates pass the test of constituting a solution to the problem of induction. This, Mackie thinks, is a mistake. Mackie has a further distinction between two kinds of necessity$_2$. Necessity $_{2.1}$ is something that would produce a deductively valid inference from C to E, and necessity $_{2.2}$ is something which grounds a judgement of probability made according to some *a priori* inductive logic, of which Carnap's or Keynes' would be examples. Mackie thinks that Hume disposes effectively of necessity$_{2.1}$ but not of necessity $_{2.2}$ and that he fails to notice that necessity$_1$ is different from either.

My criticism of Mackie on this point is that in so far as the distinction

P2614

between necessity$_1$ and necessity$_2$ is important, Hume has a better grasp of it than Mackie.

It is strange that Mackie should accuse Hume of not noticing the distinction between something's being the distinguishing feature of a causal sequence and its being the basis for a solution to the problem of induction, for Hume appears to notice just such a distinction. He considers the suggestion that if there were real powers binding cause to effect then that would justify induction. He rejects this on the ground that the fact that certain objects have possessed certain powers in the past is no reason for believing that they will in the future unless inductive method is justified (p. 91). Given that powers if they existed would constitute a distinguishing mark of causal sequences, Hume is here arguing that something which characterized causes might fail to assist in the problem of induction. It follows that he recognizes the distinction between these two tasks. Mackie does not recognize the relevance of this passage because he regards powers as constituting some form of necessity$_2$. He says that Hume's claim that powers would not solve induction rests on the assumption 'that the power needs to be inferred, that it cannot be observed' (p. 21). However, although Hume does believe that powers could not be directly observed, his argument, if sound at all, shows that they would not solve induction even if they could be perceived. If powers could be observed, it would follow that when I hold a heavy stone I can infer that if I had not been holding it and everything else had remained the same it would have fallen: powers which are known would indeed allow me to infer *a priori* what *would have* happened, *ceteris paribus*. But this tells me nothing about what the stone is likely to do if released in one second's time, or tomorrow, or whether things would have remained equal if I had released it, for without invoking the principle of induction I have no grounds for believing that it will continue to possess the power once I cease to perceive the power. Powers involve a form of *a priori* connection with their effects in just the same manner as do desires or intentions; that is, the specification of a power essentially involves reference to what it is a power to do: you cannot have a power which is not a power to produce some specific effect or range of effects, just as you cannot have a desire which is not a desire for this or that object. There is, therefore, an *a priori* or analytic connection between any power and what it is a power to do, in that reference to what it is a power *for* is essential to its identity as *that particular power*. But, as with desires, this *a priori* connection is no guarantee that the effect will occur. In so far as there is any *a priori* inference it is of the very weak form illustrated above. We can infer that if there is a power present, then if the power continues to exist and the circumstances are entirely appropriate for its actualization, then the effect must follow: just as we can infer that if a man has a desire and the circumstances for gratifying that desire are appropriate, with nothing preventing him, then, if he retains the desire, he will necessarily act on it. But no *a priori* argument can show that these conditions must, or will probably, be met in any case.

Thus powers do not solve the problem of induction whether or not they are perceptible, though they do involve *a priori* connection and a weak form of conditional *a priori* inference. This is the point on which Mackie's grasp of the distinction between necessity$_1$ and necessity$_2$ is inferior to Hume's: Hume realizes that powers do not consitute a form of necessity$_2$ and Mackie does not. Furthermore, Mackie's grasp on the distinction is flawed because he shows no sign of recognizing that anything which counted as an interesting form of necessity$_1$ would involve *a priori* connection in the way in which powers involve it, and therefore would license the corresponding weak form of *a priori* inference. It is easy to imagine uninteresting sorts of necessity$_1$. According to Hume, the tendency of the mind to associate the *relata* distinguishes causal sequences from others and therefore constitutes a form of necessity$_1$, but we are interested in finding something in the objects which will do the job. There are, therefore, at least two requirements for necessity$_1$, (i) that it be something non-psychological which distinguishes causal sequences, (ii) that it be something other than constant conjunction and contiguity as allowed by Hume. However, if the issue is to be interesting it must meet a further requirement, for Mackie regards necessity$_1$ as a form of necessity—that is as that thing which in some sense explains or constitutes the bond between cause and effect. Thus we can state as condition (iii) that any candidate for being necessity$_1$ must seem intuitively to be properly called a form of *necessity*. This condition will not inevitably be met by just anything which happened to mark off causal sequences. Suppose that only and whenever two events are causally conjoined there is the sounding of the note C sharp. Such an occurrence would mark out causal sequences from non-causal ones, as required in the definition of necessity$_1$. But such an additional element would not tell against Hume's theory. The sounding of C sharp could not plausibly be called the necessitating factor. It would merely add another constant conjunction, namely between causal links and C sharp. If we answered the question 'what is it for two events *a* and *b* to be causally conjoined?' by saying that it is for *a* and *b* to occur in succession as an instance of a law and in the presence of the sound C sharp, our answer would not be less dismissive of the concept of necessity than one which substituted reference to a tendency of the mind for the reference to C sharp. The trouble with C sharp is that it only provides another element in the conjunction, it does not lift us from the level of mere conjunctions and grant us a true necessity. However, it would seem that there could not be a third element which was a true form of necessity$_1$ which did not also involve some form of *a priori* connection. This can be

CAN WE GET RID OF EVENTS?

By IRVING THALBERG

PROFESSORS Feldman and Wierenga raise worthwhile questions (ANALYSIS, 39.1, January 1979, pp. 11–16) about the misgivings I expressed toward reductionistic programmes (ANALYSIS, 38.1, January 1978, pp. 1–9). In light of their strictures, and of some contributions to this debate which have appeared since I finished writing, I'll try to update and refine my critique. But first I should briefly elucidate a few key notions.

The crash of American Airlines DC-10, flight 191, near Chicago on 25 May, 1979, is an unimpeachable though uncommonly dramatic example of what metaphysicians nowadays mean when they speak of a particular—or concrete—event. 'Particular' contrasts with 'type' or 'kind of event'. A journalist might investigate this particular catastrophe; alternatively, she or he might do a general study of accidents of this type. As for the May 25th disaster, presumably other concrete events caused it—most immediately, the separation of an engine from the DC-10's left wing. And further concrete events resulted from the crash, beginning with an explosion and the sudden death of approximately 280 people—crew, passengers and some bystanders. Other horrified onlookers watched, and one photographed the crash. The type of event that this was cannot be close to or far from Chicago. An event-type has no temporal duration or location. It cannot result from anything, or bring about deaths and other effects. We cannot witness or take a snapshot of an event-type.

The DC-10 crash and its particular kin that I just catalogued are all changes, mostly rapid ones. But many theorists would also rank various 'unchanges' as concrete events—for instance, the weakened condition of the engine mounting from late March through 25 May. The classification of 'unchanges' is a side issue, however. What counts is that our macabre example should remind us of our rustic yet sufficiently clear and serviceable notion of a particular event.

Certainly few of us would dream of denying the existence of crashes and the like. But some philosophers appear to. Thus Professor Chisholm writes: 'there are no such things as . . . "concrete events" in addition to states of affairs' (1976, p. 115). Professor Wilson concurs to an extent, asserting: 'there is no such thing as an event distinct from a fact'; in his view, 'a fact is a true proposition', and 'an event is a truncated fact' (1974, pp. 317, 305, 304f.). Incidentally, I should mention that Professor Sellars has also disclosed that 'events are a species of [true] proposition'—though he has not argued the point (1973, p. 197). On the other hand, Professor Kim has contended in detail that 'events . . . are *exemplifications by substances of properties at a time*' (1976, p. 160, his italics). Professor Goldman agrees, emphasizing that particular actions and other events do not form 'a primitive, irreducible category'; Goldman himself 'reduces act-tokens [particular actions] to persons, act-properties, and times', and overall he believes that 'substances are primary, and events . . . derivative' (1971, p. 773). After I explore these views, I'll consider theorists who have no more use for states of affairs, facts, propositions and exemplifications than for concrete events. One frugal metaphysician of this persuasion is Professor Horgan—who at least agrees with his colleagues that 'events as particulars do not exist' (1978, p. 29).

Do the writers I quoted really mean that there are no such particular events as the May 25th tragedy? Of course not. What they advocate is a new manner of talking about the 'substance' which figures in our example—the ill-starred DC-10—that will convey all the information we have transmitted by uncouthly referring to the airplane's crash. They believe they can recreate or reconstruct our old-fashioned category of individual goings-on out of uneventful elements. You might compare their undertaking to the project of a metaphysical nominalist, who wants to do away with the abstract, mysterious attributes or Forms which Platonists ascribe to physical objects. The nominalist pledges to convey as much by talking of a melon which is round and green, as Platonists do by saying the melon 'has' or 'participates in' the Forms of roundness and greenness. However, this analogy fails, at least for Chisholm, Wilson, Kim and Goldman, since they want to manufacture concrete events out of undeniably Platonic items.

I won't look for another model of the reconstructionists' enterprise, and I must neglect countless intriguing details of the blueprints drawn up by Chisholm, Wilson, Kim and Goldman. The keystone they all work with fits my rubric of an 'event type'. Take the Chisholmian state of affairs (henceforth abbreviated 'SA', pluralized 'SAs') which is Brutus killing Caesar. According to Chisholm, SAs like this 'are abstract entities which exist necessarily and . . . are such that some but not all of them occur' (1976, p. 114). The SA of Brutus killing Caesar exists at all times, even when Brutus is not engaged in mayhem. The SA would exist if Caesar died a natural death, or if he and Brutus had never been born. For Chisholm says that SAs 'are in no way dependent for their being upon the being of concrete, individual things. Even if there were no concrete, individual things, there would be indefinitely many SAs' (p. 114).

SAs alone, however, do not give you a particular event. Nor is it enough when concrete, individual things like Brutus and Caesar abound. The SA must 'occur' as well. But what is it for SAs to 'occur'? Here Chisholm fills a gap in his previous writings. He explains that properties are involved in SAs: 'that SA which is Smith being ill', for example, 'implies Smith to have the property of being ill'; and more important, the SA becomes 'concretized' when Smith acquires 'the property of being ill' (p. 125). Time also enters the picture when Chisholm says that the 'SA which is *Brutus killing Caesar* was concretized in 44 BC by . . . Brutus and Caesar' (p. 125, original italics). Thus, for SAs, to occur is to be concretized; and they achieve concretization when a substance comes to possess a property at a time.

Wilson's reconstruction of a particular event from a fact turns out similarly. He announces that a 'fact is constituted by an individual [or set of individuals], a property and a time' (1974, p. 311). Analysing 'the fact that Columbus discovered America in 1492', he remarks: 'if you . . . subtract one or more constituents of a fact, then, if all goes well, what you are left with is an event . . . If we leave out the agent we can get the discovery of America in 1492' (*ibid.*). Thus we reach Wilson's theory that 'an event is a truncated fact'.

I would like to investigate concretization, truncation—and especially the startling dictum of Wilson and Sellars that the particular goings-on of this world are identical with a sub-class of the true propositions which humans can assert and believe. I am sorely tempted to inquire how the crash, which happened near Chicago on 25 May, but is no longer taking place, could be equated with any of the following: (i) the fact that flight 191 went down near Chicago on 25 May—a fact which never was near Chicago, and which is still a fact today; (ii) the proposition that flight 191 crashed, which was true everywhere, not just near Chicago, and which remains as true now as it was on 25 May; (iii) some truncated version of (i). However, I drop these questions about Chisholm's and Wilson's ontologies, because it should have become plain that their basic materials—substances, properties, times—and at least Chisholm's fundamental procedure, concretization, are virtually the same as those favoured by Kim and Goldman.

Clearly, properties are the main ingredients in Kim's and Goldman's recipe for particular happenings, and both writers offer us much information on them. But for our limited purpose you can characterize them as I described event-types. Indeed, Kim calls properties 'generic events' (1976, pp. 162ff.). At any rate, a generic event, or event-type, does not a particular make. Not even if you throw in substances and temporal instants or intervals. Just as SAs have to become concretized, a property must be displayed by a substance at the right moment.

In my earlier discussion, I grumbled that the occurring of SAs and the exhibiting of properties sounded eventful. Since than I came across a more recent essay of Kim's where he mentions 'the *event* of x's having [property] *P* at time *t*' (1977, p. 103). A lapse, no doubt, but a revealing one. I still suspect that elusive but unavoidable particular goings-on have been smuggled into the analysis which was designed to replace our primitive notion of a concrete episode.

Chisholm may have anticipated this sort of objection. Anyway, a counter-argument emerges from his treatment of a somewhat rococo example: 'The third occurrence of [the SA of] Jones being robbed contributed causally to his illness'. Chisholm's translation and commentary read:

> 'There is a time *t* such that: Jones being robbed occurs at *t*; Jones being robbed is such that at *t* it has been concretized exactly three times; and Jones being robbed contributes at *t* to Jones being ill' . . . [T]ruths like [this] ostensibly pertain to the particular occurrences of SAs . . . But . . . each . . . can be construed as pertaining just to individual things, properties, and SAs (1976, pp. 135f.).

Feldman and Wierenga (henceforth abbreviated 'F&W') in any case reply to my complaint along similar lines. F&W contend:

> we have been given no reason to think that if a SA [occurs] then there is a concrete event which is the occurrence of that SA . . . Chisholm's theory says that there are no such things. True sentences purporting to involve reference to particular occurrences . . . are provided with paraphrases that are about SAs and their properties (p. 13).

Formalistically speaking, Chisholm and F&W are on safe ground. A reconstructionist only has to avoid quantifying over 'occurrences', 'concretizations' and 'exemplifications'. He can refuse to say: 'There is an occurrence of the SA of Jones being robbed (an exemplification by Jones of the property of being held up)'. He can delete these compromising singular nouns, and insert the verbs 'occur' and 'exemplify', altering syntax as required. Then he is not formally committed to acknowledging 'particular occurrences' of SAs, or concrete exemplifications of properties. He only reifies SAs which occur, and properties which are exemplified. But doesn't that leave a dense fog around the reconstructionist's reports of SAs becoming concretized and properties getting exemplified?

Naturally I would beg the question against a reconstructionist if I maintained that his statements only make sense if we interpret them as recording encounters between a substance and a property or SA. Still, the reconstructionist has undertaken to convince us, presumably by reasoning, that a garden-variety statement like 'Jones is robbed at *t*' can be true without having a stickup to match it. We are told that this is really equivalent to saying 'Jones concretizes (or exemplifies) the event-type of being robbed'. Yet how can this strategy work if the reconstructionist's paraphrase creates the same problem all over again? For isn't there an episode of concretization or exemplification unfolding amidst his stage machinery of event-types? It would be dogmatic if reconstructionists proclaimed without ado that 'there are no such things as' particular robberies and crashes. Would it be any less arrogant if they now lay down the convention that 'there are no such things' as particular concretizations and exemplifications?

I won't speculate on that, but I'll turn to another hunch I ventilated in my previous essay. I suspected more eventfulness when reconstructionists append a temporal specification to their reports of SAs occurring and properties becoming exemplified. I wondered if prepositional phrases like 'at (during, prior to, after) time *t*' imply a backdrop of unique particular events—the creation of our universe, the miraculous birth of Jesus—or particular landmark happenings which fall into a regular pattern, as the planetary motions do. To say that the ill-fated DC-10 went down in 1979 would be to say that it crashed 1979 years later than the alleged miraculous birth. Years, in turn, would be measured against concrete stellar and seasonal changes.

F&W are unperturbed. They seem to admit that our workaday 'dating systems contain expressions that make reference to events'; but they think this 'in no way conflicts with the claim that events are triples of substances, properties, and times' (p. 15)—the latter being nailed down by reference to further events. F&W also attempt to quash my charge of circularity. F&W propose to

> fix the referent of the expression 'Time 1' by saying that it shall refer to now (the present moment). A short time [*sic*] later we might establish the referent of 'Time 2' in a similar way. A unit of time would then be the interval between Time 1 and Time 2. Such a system might be awkward to use unless a unit of time coincided [*sic*] with some regular and natural occurrence . . . [But] such a connection . . . does not seem to be a logical requirement for reference to times. So it is implausible to think that we can't refer to times without somehow referring to events (p. 16).

F&W's putative counter-example only reinforces my qualms· F&W's baptismal procedure runs into difficultly with a standard analysis of their first crucial term, 'now'. Reichenbach's familiar 'token-reflexive' theory of tensed discourse (see his 1947, pp. 248–87) equates 'now' with 'the time at which *this token* is uttered'; and the token he is alluding to is the sequence of noises, marks, sign-language gestures or whatever that you have produced, which contains a token of 'now'. Perhaps we should illustrate: A flight controller was broadcasting while the DC-10 crashed. If he had made the sounds, 'Flight 191 is exploding now!', then he would have pinpointed a moment through his speech-act of uttering the token with 'now'. Likewise when F&W intone a token of ' "Time 1" . . . shall refer to now'. So aren't they labelling Time 1 by means of a simultaneous event, their ceremonial act of naming Time 1? Obviously F&W do not have to endorse a token-reflexive analysis of 'now'. But until they put forth an eventless rival theory, I shall continue to suspect that 'we can't refer to times without somehow referring to events'. Moreover, I see no way to reconstruct events without times.

So much for F&W. The reconstructions they defend all feature event-types and mysterious concretizing or exemplifying relationships between a substance and its event-type. I said I would discuss other metaphysicians who plan to get rid of concrete events while not introducing event-types. I quoted Horgan's proclamation, 'events as particulars do not exist'. Rather than conjure up non-particulars, Horgan tries statements, specifically reports of action. He imagines that Jones has apologized to someone by uttering suitable words. Many philosophers would assume that one, perhaps two, social events have taken place. Horgan demurs:

> we need not assume that there exists an entity (or two . . .) answering to the descriptions 'Jones's saying "I apologize" ' and 'Jones's apologizing'. Rather, we need only recognize that the following statement is true: 'Jones said "I apologize", and thereby he apologized' . . . [We] can dispense with the notion of redescribing an action, and also with actions themselves (1978, p. 37).

One worry I have about this is an offshoot of my reaction to F&W's christening of Time 1. I wonder: doesn't a statement have to be stated, and doesn't a speech-act take place when you state what Jones did? My remaining uncertainty is directed to another theorist besides Horgan. Like Horgan, Professor Aune stays away from both particular events and event-types. He too dwells on 'the supposed acts that agents perform', which he thinks may be supernumerary; but Aune would

> employ predicates where an action theory employs singular referring terms. Thus, instead of [postulating] John's act of hitting Tom . . . [we] might say . . . that . . . John hit Tom . . . [S]ince predicates [like 'hit Tom'] are not referring terms, they do not commit [us] to actions or events (1977, p. 44).

Regarding Aune's 'predicates' as well as Horgan's true statement about Jones, my doubts are embarrassingly simple-minded. Could predicates and true statements be all that we are concerned with when we gossip about Jones and John? Could predicates and true statements serve as ultimate building blocks in a reconstruction of our antiquated category of particular goings-on? Can't we penetrate beyond the true statement, 'Jones said "I apologize", and thereby he apologized'? Why, after all, is the statement true? Horgan should at least tell us why we provide no illumination if we say his statement is true inasmuch as it tallies with Jones's linguistic and social antics. The same for Aune. Why is it correct to employ the predicate 'hit Tom' rather than 'hugged Tom'? Is 'hit' just the right predicate, and 'hugged' the wrong one? Or might there be a punch but no embrace in the offing?

I cannot resolve these perplexities here. It has scarcely been my goal to demonstrate that amateurs of event-types must inevitably bring in contraband particular goings-on, camouflaged as exemplifications and times. Nor have I tried to prove that if you limit yourself to true statements, or to predicates, you necessarily get entangled with particular speech episodes and events that guarantee truth or correct predication. The *onus probandi* is on reconstructionists and their defenders, including F&W. At this juncture, we have imposing reasons to doubt that analytical metaphysics can get by without concrete events.

University of Illinois at Chicago Circle © IRVING THALBERG 1980

REFERENCES

[1] Aune, B. (1977) *Action and Reason.* Dordrecht: Reidel.

[2] Chisholm, R. M. (1976) *Person and Object.* London: Allen & Unwin.

[3] Goldman, Alvin (1971) 'The Individuation of Action', *Journal of Philosophy,* LXVIII, pp. 761–74.

[4] Horgan, T. (1978) 'The Case Against Events', *Philosophical Review,* LXXXIII, pp. 28–47.

[5] Kim, J. (1976) 'Events as Property Exemplifications', in M. Brand and D. Walson, eds., *Action Theory.* Dordrecht: Reidel, pp. 159–77.

[6] ——, (1977) 'Causation, Emphasis, and Events', *Midwest Studies in Philosophy,* II, pp. 99–103.

[7] Reichenbach, H. (1947) *Elements of Symbolic Logic.* N.Y.: Macmillan.

[8] Sellars, W. (1973) 'Actions and Events', *Nous,* VII, pp. 174–202.

[9] Wilson, N. (1974) 'Facts, Events, and their Identity Conditions', *Philosophical Studies,* XXV, pp. 303–21.

ACTIONS AND BODILY MOVEMENTS: ANOTHER MOVE

By DAVID W. D. OWEN

I

IN 'Actions and Bodily Movements' (ANALYSIS 38.3, June 1978, pp. 137–40) James Montmarquet discusses a problem that arises for what appears to be a basic assumption of many modern theories of action, especially that of Donald Davidson. When I raise my arm, volitionists like Prichard identify the action with a certain mental event that causes my arm to rise.[1] According to Montmarquet, theorists such as Davidson identify the action of arm raising, not with a prior event that causes my arm to rise, but with the event that is my arm rising itself. Montmarquet then goes on to argue that if my *action* of arm raising is the bringing about of the bodily movement that is the *event* of my arm going up, then there is something wrong with identifying the action (the bringing about of an event) with the very event that is brought about. I agree with Montmarquet that there is this problem for anyone who thinks that actions are bodily movements. But I wish to argue that there is more to it than Montmarquet uncovers. And when the issues are uncovered, it is no longer obvious that it is right to claim that Davidson holds that the action of my raising my arm is identical with the event that is my arm rising.

II

Suppose I move my hand, thereby pushing a button which causes a mechanism to be activated which in turn causes a flag to be lowered. We have several events here: the movement of my hand, the button being depressed, the mechanism being activated and the flag being lowered. But according to Davidson there is only one action, which can be variously described in terms of any of these events: moving my hand, pushing the button, activating the mechanism, or lowering the flag. In the passage which Montmarquet cites, Davidson says,

> We must conclude, perhaps with a shock of surprise, that our primitive actions, the ones we do not do by doing something else, mere movements of the body—these are all the actions there are. We never do more than move our bodies—the rest is up to nature. ('Agency', p. 23)

Even when an action is described in terms of an event, such as the flag being lowered, that occurs far outside the confines of my body, there is always another description of that very same action in terms of an event that occurs closer to home. That is, we can always describe an action in terms of an event that is a 'mere bodily movement'. It is clear that Davidson believes that the one action in our example is not identical with the event that is the flag being lowered, nor is it identical with the event that is the button being depressed nor that of the mechanism being activated. But does it follow, as Montmarquet supposes, that he believes that it is identical with the event that is my body moving?

---

[1] "When I move my hand, the movement of my hand, though an effect of my action, is not itself an action, and no one who considered the matter would say it was, any more than he would say that the death of Caesar, as distinct from his murder, was an action or even part of an action." "Acting, Willing, Desiring", H. A. Prichard, *Moral Obligation* (Oxford, 1968), p. 191.

In order to answer this question, we must take note of an important distinction. The phrase 'bodily movement' is ambiguous. It can be thought of as the nominalization of either of the following two sentences:

> I moved my body.
> My body moved.

Let us call what corresponds to the nominalization of the former sentence a bodily movement-a, and what corresponds to the nominalization of the latter sentence a bodily movement-e. A bodily movement-e is the 'mere' event of my body moving. A bodily movement-a is the action of my moving my body. When I move my body, my body moves. But it does not follow that my action of my moving my body is identical with the event of my body moving. It is not necessarily any more true to say that a bodily movement-a is identical with a bodily movement-e than it is to say that an action of lowering a flag is identical with an event of a flag being lowered.

Davidson believes that we never do more than move our bodies. All our actions can be described as bodily movements-a. Does he also believe that all our actions are identical with the events that are bodily movements-e, i.e. with the events in terms of which they are described as bodily movements-a?

III

The passage from 'Agency' already quoted does not help to answer this question. 'We never do more than move our bodies' only commits Davidson to the view that all actions can be described as bodily movements-a. And the claim 'mere movements of the body' are all the actions there are is ambiguous: it may mean that all actions are identical with events that are bodily movements-e, but it may simply mean that all actions can be described in terms of those events, i.e. as bodily movements-a. Montmarquet cites, but does not quote, another passage in favour of his interpretation:

> if a man's arm goes up, the event takes place in the space-time zone occupied by the arm; but if a man raises his arm, doesn't the event fill

> the zone occupied by the whole man? Yet the events may be identical. ('The Logical Form of Action Sentences', p. 116)

This passage occurs in a reply to a comment on the original paper. Furthermore, it only says that the events *may* be identical. It is hardly decisive. And in another paper Davidson criticizes Melden for holding just the view that Montmarquet wants to attribute to Davidson himself:

> Obviously, the problem is greatly aggravated if we assume, as Melden does (*Free Action* p. 85) that an action ('raising one's arm') can be identical with a bodily movement ('one's arm going up')[2]

Finally, the passage from 'Agency' that Montmarquet quotes in full, which contains the insight that Montmarquet develops into a criticism of the view that actions are bodily movements-e, is surely evidence that Davidson does not hold that view:

> A man who raises his arm both intends to do with his body whatever is needed for his arm to go up and knows that he is doing so. And of course the cerebral events and movements of the muscles are just what is needed. So, though the agent may not know the names or locations of the relevant muscles, nor even know he has a brain, what he makes happen in his brain and muscles when he moves his arm is, under one natural description, something he intends and knows about (p. 12).

The 'natural description' under which this action is intentional is, presumably, 'doing with his body whatever is needed for his arm to go up'. This action is identical with the action described as a bodily movement-a, i.e. 'raising his arm', but it doesn't follow that the action is identical with the event that is the arm going up. If this passage led one to identify the action with any one event, it would be more plausible to pick a cerebral event or the contraction of the muscles than it would be to pick the movement of the arm.

IV

All actions are bodily movements-a. This is all that Davidson is committed to. As we have said, it is clear that actions are not to be identified with events, such as the flag being lowered, that occur outside the confines of the agent's body. We have seen that it is at best dubious that Davidson holds that all actions are bodily movements-e, and in any case Montmarquet has argued that there are grave problems with that position. Are we then to hold that actions are the events that are muscle contractions or neurons firing? It was held that it was no more true to identify the action of moving my body with the event, i.e. the bodily movement-e, that was brought about by the action than it was to identify the action of lowering the flag with the event, i.e. the flag being lowered, that was brought about by that action. But then doesn't it follow that it is just as implausible to identify the action of bringing about neurons firing (or muscle contractions) with the event brought about by that action, i.e. with the event that is the neurons firing (or the muscles contracting)? The problem is quite general: it doesn't seem plausible to identify the action of bringing about some event with the very event that is brought about. The problem arises whether the event brought about is a flag lowering, an arm rising or a muscle contraction. Perhaps the beginning of wisdom is to cease to look for the one event with which an action is identical. It may be that actions are not events brought about, but the bringing about of events. To say that an action is the bringing about of an event is not to say that it is some other event that causes the event brought about. For then the same question would arise: is the action of bringing about that other event identical with that event brought about? We are then left with the question what actions *are* and how they are related to the events brought about by them. But that is a question that neither Montmarquet, Davidson nor I have attempted to answer.[3]

---

[2] 'Actions, Reasons and Causes', *The Journal of Philosophy* LX (1963) reprinted in *Readings in the Theory of Action*, ed. N. S. Care and C. Landsman (Indiana, 1968), p. 197.

Corpus Christi College, Oxford © David W. D. Owen 1980

---

[3] I have benefited much from discussions about the substance of this note with J. Hopkins, J. Hornsby and J. Annas.

KILLING AND AVERAGE UTILITY

By Michael Neumann

RICHARD Henson, in an article entitled 'Utilitarianism and the wrongness of killing' (*Philosophical Review* 80, 1971, pp. 320–7), asserts that the principle of average utility can lead to "paneuthanasia". Painless killing of those below the average level of happiness in a population, he says, may raise the average, and could therefore be obligatory. Suppose, for example, we can only kill Luke painlessly, or refrain from such action. His death will have no effect on others. Along with John and Mary, he makes up a population whose (roughly constant) levels of utility are as follows:

John	10
Mary	9
Luke	8.

By killing Luke, we can raise average utility from 9 to 9½, so we must do so. If we get a similar chance to kill Mary, we must do so as well: we would thereby raise the average level of utility from 9½ to 10.

Henson's argument relies on a failure to define the average utility of an act. He apparently would have us define it as the total utility of the act divided by the number of persons who exist after the act is performed. He would, in other words, have us conceive of the situation as follows:

	Initial utility levels	Utility levels after the 1st killing	Utility levels after the 2nd killing
John	10	10	10
Mary	9	9	—
Luke	8	—	—

Luke and Mary do not get counted into the average after they are killed, because they do not exist.

Average utilitarians do not, however, calculate in this fashion. They define the average utility of an act as its total utility divided by the number of those who possess at least one of two qualifications: (i) they exist when the act is performed, (ii) they exist after the act is performed. Where Henson would have us put dashes, signifying that there *is* no one to be counted into the average, average utilitarians would put zeros, signifying that the decreased *do* get counted: they existed at the time the act was performed.[1] This changes the average utility of the first killing from 9½ to 6⅓, a decrease from the initial average utility level of 9. The average utility of the second killing drops from 10 to 3⅓. Both acts are therefore wrong, and Henson's problem dissolves.

Trent University © MICHAEL NEUMANN 1980

---

[1] This method of calculation should be understood to measure utility on a scale which, like Henson's, includes negative values. Otherwise there is no fate worse than death: others' utility being equal, to prolong even the most painful existence is never worse than to end it.

PHENOMENOLOGICAL SPECIMENISM

By J. M. HINTON

I

'PHENOMENOLOGICAL Specimenism' is a name for a belief I see no reason to share. There isn't, as far as I know, the quite particular way of neutrally saying exactly what you underwent, that it says there is. The phenomenological, conceived in the Specimenist way, which is Anglo-Saxon whatever else it may be, is as 'absent' as God is for Heidegger.

Phenomenological Specimenism says that some neutral experience-reports satisfy the following condition.

Condition PS. That of having enough precision or exactitude, or whatever else it might take, to support at least one of the following pair of philosophical ideas substantively, i.e. in a way that is not vacuous or jokey:

Idea A That among the particular events of which the given human subject was the conscious subject, on the occasion when his neutral experience-report was true, there was one specimen event of a species whose occurrence was explicitly stated by the report, which specimen could have occurred with him as its conscious subject if the case had been whichever it in fact was not: illusion if in fact it was perception, perception if in fact it was illusion. (Idea A can best be regarded as just not applicable, and therefore not put forward, in those cases of illusion where there could not be such a thing as the corresponding perception, such as the cases of the so-called impossible figures.)

Idea B That where the case in which the neutral experience-report is made is in fact one of perception, the following thing is true. If the subject later gets the perfect illusion of that very same perception, then there will have occurred on those two occasions two specimen events, the second 'qualitatively' indistinguishable by the subject from the first, patently of the same species. The neutral experience-report explicitly stated the occurrence of an event of that species.

II

Postponing to a third section a certain reason for thinking that Phenomenological Specimenism is a widely held belief, I want next to make five comments on the above.

(i) A report that satisfied Condition PS would thereby, I believe, satisfy what I will call Precondition P: that it, the report, be distinguishable in a way that matters from what I have elsewhere[1] called the nearest perception-illusion disjunction—when there is such a thing. For I think it would be difficult to claim that a perception-illusion disjunction could satisfy Condition PS. A plausible view would be that it could support Idea B in a jokey way and Idea A not at all. Awkwardly however, I do not have space, here, to say exactly what a perception-illusion disjunction is and *can* do. Or to go accurately into other questions that arise about that kind of sentence, statement and proposition. Quizzed on that, I would refer the reader to the work cited, much of which is about that. (And indeed other terms and assumptions that I have to use and make here might usefully be seen in that context.) But roughly speaking and for example, suppose that what the neutral experience-report says is that I am having the experience 'as of' seeing a dagger. Then the nearest perception-illusion disjunction says that I either visually perceive a dagger or am getting the illusion of doing so. That there is little difference between the report and the disjunction is a contention of the book to which I have referred. Any impression that it is, philosophically, *un*contentious should be dispelled by reflecting that it says the report does not satisfy Condition PS. (The first paragraph of III, below, can also be consulted in that regard.)

(ii) Although I contend that no neutral experience report satisfies Condition PS, I do not undertake to prove that. I am not adopting the method of pointing to something that I can demonstrate to be a mistake, and suggesting or showing that it is common to a certain class of doctrines. Instead I am just taking a stand until overwhelmed by contrary reasoning. (Incidentally, this was also the case in *Experiences*, though one reviewer thought otherwise.) The method I am not adopting may seem superior to the method I am adopting. And certainly we are helped when we are shown mistakes that can be involved in doctrines of a certain sort. Yet even then we may still need a variety of particular objections or obstacles to particular ways in which particular doctrines get recommended—if the obstacles block dubious doctrines that do not turn out to commit the given mistake or mistakes. So perhaps there is something to be said, as well, for a kind of fundamental objection or alienated position that, far from even appearing to supersede such particular objections or obstacles, manifestly requires to be defended against all comers by a fine collection of them. And that is the kind of fundamental objection I am endorsing, unsatisfying in some ways though it may be.

(iii) The objection is to a class of doctrines that goes beyond the class of 'sense datum' doctrines, and yet leaves some 'sense datum' doctrines out. The class includes doctrines that commit no particular mistake discriminated and exposed either by Anscombe in 'The Intentionality of Sensation: A Grammatical Feature' (*Analytical Philosophy*, ed. Butler, 2nd series, Blackwell 1965) or by Quinton in 'The Problem of Perception' (*Mind* 1955, pp. 28–51; reprinted in *Philosophy of Perception*, ed. Warnock, OUP 1967). For its members need not posit sense-data conceived as sensible objects, nor need they posit any kind of inference from experience to external objects—they can say with Quinton that the relation between experience and beliefs about external objects is causal only. The stand is of course like Austin's. As just intimated, it is not confined to *sensa*—or, one may add, *sensibilia*.

---

[1] *Experiences,* Clarendon Library of Logic and Philosophy, OUP 1973 (v.s.v.p. 92)

(iv) Certain physical models seem appropriate to what I have called Ideas A and B, above. To idea A there corresponds the model of a certain copy of a certain book, which stands in a certain place in the bookshelf but could have had different neighbours there. To idea B there corresponds the model of two different copies of the same book, your copy and mine say, where if handed yours I cannot tell that it is not mine though the converse may not be true. However, everything might slip away if one said that it was the appropriateness of those models that was in question. What is in question is whether neutral experience-reports can have properties such that their having them supports Idea A or Idea B in the required, non-vacuous and non-jokey, way.

(v) I have deliberately tried to avoid stating Ideas A and B in such a way that they involve demonstrable mistakes. Like the mistake of forgetting the impossible figures and assuming that to every illusion there corresponds a possible perception. Or the mistake of combining an unadvertised and tacit 'What it is' with an avowal of *absolute* self-restriction to how things seem. Or the mistake of thinking that 'Given X, you cannot tell that it is not Y' entails 'Given Y, you cannot tell that it is not X'. Rather than being demonstrably wrong, Ideas A and B seem to me in the end to turn out, through lack of *effective* support, to be typical cases of philosophically beating the air, where you find yourself trying to prove something by emphatic emission. That is often, when they cease to be covert, the status of normally covert ideological assumptions by which one's conception of the fundamental, the substantive, the explanatory, the clear, the plausible, the humorous and the otherwise canonical is determined.

III

Phenomenological Specimenism is understandable, unsurprising, in one who finds that he can accept certain doctrines of sense-data in spite of all the obstacles that can be erected to particular ways in which particular doctrines of that sort get recommended, and in spite of the various mistakes that have been alleged to be common to classes of doctrine that at least overlap with them. Ways of trying to support Phenomenological Specimenism: that is exactly what certain, indeed most, sense-datum-isms are. The sense-datum-statement is a special kind of neutral experience-report then. The discovery, of there being so much as one true sense-datum statement as such statements have *usually* been conceived, could be called Proof PS/s, the sense-datumist proof of Phenomenological Specimenism. The satisfaction of Condition PS and Precondition P would tag along behind. However, I now come as promised to give a reason for thinking that Phenomenological Specimenism is a widespread belief: and the reason happens also to be a reason for thinking that the belief is, more surprisingly, alive and well in many people who do not go in for sense-data at all. PS without PS/s.

I shall make the following assumption. I shall assume that it will be sufficient, in order to give a reason though not a conclusive reason for thinking that Phenomenological Specimenism is a widespread belief, to show that Phenomenological Specimenism could very well be what underlies a certain type of mainly informal philosophical discussion that is almost certain to be going on somewhere in the English-speaking world at any given moment, in the main without undue solemnity but with a degree of seriousness that is not always easy to gauge. A notably un-solemn discussion of the kind I mean at the Oxford Philosophical Society, involving Quinton and Lucas among others, lives in memory as an example even if no example need strictly be given. The extremely familiar kind of discussion I have in mind concerns various combinations of subjective experiences 'as of' this and that, combinations which as far as anyone knows are purely hypothetical. For example the combination of the visual experience, as of seeing a cube in front of one, with the tactile experience, as of feeling in the same place a sphere. More complex and interesting combinations get mentioned. The question, in that kind of discussion, is as to what judgements it might be best for us all to make about our shared situation if the combinations began and continued to occur in a massively and systematically shared way.

In trying to show that Phenomenological Specimenism could very well be what underlies such a discussion I shall further crucially assume the following thing. That, in that kind of discussion, when it is serious, not only each hypothetical combination of experiences, but also each supposition as to the sharing of the experiences is being assumed to represent a *possibility*, in some sense of the term whose content is adequate for the context. For the purpose of making 'suppositions for the sake of argument' that one was not prepared to characterise as representing possibilities would be obscure under those conditions.

Given, then, that the suppositions are being assumed to represent possibilities, we can ask, and we can try for a determinate concept of possibility for the context by asking, why. Why is that being assumed, on what ground is it being assumed, what underpins that assumption? The answer could very well be that Phenomenological Specimenism underlies and supposedly underpins it. For the underlying thought might be this: that distinct specimens *might* recombine in any given way unless there is some *known* reason why they cannot. That is perhaps a semi-epistemic modification of an idea originally ontological and atomistic: indivisible things can recombine in any way. It certainly suggests Phenomenological Specimenism as above defined. The question really is what underpins those discussions if not Phenomenological Specimenism.

I have now tried to show what I undertook to show, and I end at a tangent as follows. One can hardly prove that Phenomenological Specimenism is not involved in one of those discussions by just asserting that each supposition 'is at least a logical possibility'. For (a) that is unargued, or in other words how do you know? (b) It could itself rest on Phenomenological Specimenism. (c) The absence of an internal contradiction in a supposition is conventionally sufficient for a logical possibility but not for a 'might be' or a 'might have been', whereas a possibility in those discussions had better at least be one of these. (d) The very locutions, 'logically possible' and 'a logical possibility', are suspect, as if sprung from a dubious inversion whereby 'Not: logically, not possible' becomes 'Logically, not not possible' and so 'Logically possible'. (It does not go over very well if you substitute 'provable' for 'possible'.) Neither by such a derivation nor in any other way can anyone be compelled to regard the absence of logical falsity as constituting any kind of a possibility.

Worcester College, Oxford

© J. M. Hinton 1980

CHISHOLM'S DEFINITION OF THE EVIDENT

By WAYNE WASSERMAN

THE definition of being evident which Roderick Chisholm offers in *Theory of Knowledge* (Prentice Hall, Englewood Cliffs, 1977) gives rise to two unacceptable results: (E) knowledge comes in only two degrees of reasonableness and (C) nothing we know about the external world is more reasonable than anything else we know about it. To see this consider some of his key definitions and claims.

Chisholm uses *accepting a proposition* and *more reasonable* (which he holds to be equivalent to *epistemically preferable, more respectable epistemically,* and, at least this is suggested by his comments on pages 5 and 6, *better justified*) as basic undefined concepts. Using these he defines *withholding a proposition p* (page 6) as *accepting neither p nor not-p,* and *beyond reasonable doubt* (page 7) as *more reasonable to accept than to withhold.* He then defines *certain* (page 10) and *evident* (page 12):

D 1.4 *h* is certain = df *h* is beyond reasonable doubt & there is no *p* more reasonable than *h*.

D 1.5 *h* is evident = df *h* is beyond reasonable doubt & for all *p*, if *p* is more reasonable than *h*, then *p* is certain.

(Chisholm relativizes these definitions to persons. Nothing turns on my simplification).

These definitions logically entail that all certain propositions are evident. For, if *h* is certain then it is beyond reasonable doubt, and, since nothing is more reasonable, it is trivially true that all *p* more reasonable are certain. The definitions also logically entail that (1) all non-certain evident claims are equally reasonable and (2) all certain propositions are equally reasonable. The proof of (1) is this. Suppose *p* and *q* are non-certain and evident. If *q* were more reasonable than *p*, then *p* could not be evident, for there would be a non-certain claim more reasonable than *p*. The proof of (2) is similar. In light of the above it follows that (3) evident propositions come in only two degrees of reasonableness.

When (3) is combined with Chisholm's requirement that everything known is evident (see D 6.4 page 110) we get (E) that knowledge comes in only two degrees of reasonableness. Chisholm also seems to hold (pages 11 and 12) that nothing about the external world is certain, hence that all we know about the external world is non-certain and evident. By (1) it follows that (C) nothing known about the external world is more reasonable than anything else known about it.

When I accept Chisholm's (page 16) commonsense starting point, that we know 'pretty much that which, upon reflection, we think we know,' I conclude that these results are false. Let me indicate in (a)–(e) a few of my reasons for that claim.

(a) Consider some propositions, each more reasonable and reliable than those following it, I knew one night: (*p*) I was appeared to stereo-ly; (*q*) something near me looked and sounded like a stereo; (*r*) a stereo which looked and sounded just like mine was there; (*s*) Linda (who had retired earlier, closed the door and not reappeared) was in the bedroom of our second floor apartment. According to Chisholm, only *p* would be certain. Thus *p*, *q*, *r*, and *s* are known propositions of more than two degrees of reasonableness, and *q*, *r*, and *s* are about the external world and differ in epistemic respectability.

(b) We know propositions which differ in probability, but which are essentially similar in terms of informational content. Consider two urns. In one case the odds of picking black are five hundred million to one, in the other three hundred million to one. Presumably I can know that balls picked at random from both urns are black, although the probability is greater in one case than in the other. The statistical nature of many key and highly confirmed scientific laws (not only from quantum theory but also from thermodynamics and the theory of gases) suggests related examples for those who dislike urn examples. It thus seems to me that even if reasonableness is not determined solely by a claim's probability, but is somehow the product of a claim's probability and its amount of informational content, we have claims which differ in reasonableness, but which are known and are about the external world. (This view of reasonableness has been suggested by Keith Lehrer in several places, for example in 'Reason and Consistency,' in *Analysis and Metaphysics,* ed. Keith Lehrer, D. Reidel, Dordrecht-Holland, 1975).

(c) Suppose *p* and *q* are non-certain, each trivially implies a non-certain claim *r* (e.g. $p \vee q$), and *S* knows both and sees the two implications. Claim *r*, in general, will be more reasonable than either one of *p* and *q*. For *r* to be false both *p* and *q* must be false "at the same time."

(d) On the most common view of science the super-structure of scientific knowledge is based on observations. On this view the observations may be rejected in extreme cases, but in general they are more reliable, more secure, less open to revision and better justified than the super-structure of knowledge which they support.

(e) Consider an example due to Robert Audi. Suppose *W* has done just enough research to add *p* to our body of knowledge. Suppose *Y* (perhaps *W* at a later time) does all *W*'s research and conducts independently confirming tests. Both know *p*, but *p* is more reasonable for *Y* than for *W*. In general it seems that knowledge based on hypothetico-deductive and probabilistic methods can always become better confirmed, and such knowledge varies almost continuously in justification.

(a)–(e) each contain reflections about the relative epistemic status of things we know which support the denials of (E) and (C). Taken together they seem overwhelming. Weakening Chisholm's claim that nothing about the external world is certain does not seem to help. Suppose some simple claims (e.g. Moore's claim to see his own hand) are certain. Chisholm would still be committed to the claim that except for some simple claims which are just as reasonable as the *cogito*, all knowledge of the external world is equally reasonable.

It would be more plausible to agree that (C) and (E) are false when 'more reasonable' is used with its ordinary sense, but hold that when the expression (and related terms) are used as Chisholm does the claims are true. Showing the claims are false in their ordinary senses would thus not show that Chisholm's typographically similar claims were false.

I cannot accept this suggestion. First, if the falsity of the claims in their ordinary senses does not entail their falsity in the senses Chisholm uses, then Chisholm's notion of reasonableness differs considerably from the ordinary notion. Since the notion is left undefined it becomes unclear what Chisholm is talking about. This is important, for claims made in *Theory of Knowledge* which had seemed clear and plausible to the reader, who had mistakenly taken 'more reasonable' to have pretty much its ordinary sense, would become obscure and be of indeterminate truth value. Second, this move requires that my intuitions are not generally reliable indicators of epistemic status, and this seems contrary to Chisholm's commonsense starting point.

The particular account of being evident which Chisholm gives plays no substantial role in *Theory of Knowledge*. It is required neither by the insight that being known entails being evident, nor by Chisholm's general foundationalist picture. I do not see any reason for accepting it. On the other hand, it gives rise to unargued for results which seem upon considered reflection to be false. I suggest that at the very least it is more reasonable to withhold than to accept this account of the evident.[1]

University of Nebraska-Lincoln
© WAYNE WASSERMAN 1980

---

[1] I wish to thank Charles Sayward, Philip Hugly, Mark Timmons, and especially Eric Russert Kraemer.

GOLDMAN ON EPISTEMIC CONJUNCTION

By FRED CHERNOFF

IN a recent paper[1] Alan Goldman argues that the principle of conjunction is invalid when applied to propositions known to be true. He states that his position conforms with that of Henry Kyburg. A source of motivation for their view is the lottery paradox, which arises as follows. Consider a rule of acceptance for justified belief, according to which h is accepted whenever prob$(h, e) > m/n$, where e is our total evidence, m and n are integers and $0 < m/n < 1$. We hypothesize a fair lottery in which one and only one ticket wins. Let there be $n+1$ tickets. If we ask whether ticket 1 will win, our rule of acceptance answers in the negative. The probability that it will not win is $n/n+1$, which is greater than m/n. If we ask whether ticket 2 will win, our rule of acceptance furnishes a negative reply in the same way. Similarly we must accept that ticket 3 will not win, . . . , and that ticket $n+1$ will not win. When we conjoin these results we see that ticket 1 will not win, and ticket 2 will not win, and . . . , and ticket $n+1$ will not win. Therefore no ticket will win. But this contradicts our assumption that the lottery is fair.

There are several moves that one may make to avoid this evil result. They consist of denying one or more of the principles that warrant the inferences required to state the paradox. Goldman chooses to reject the principle that sanctions the conjunction of several propositions individually derived by our rule of inference. Moreover, Goldman claims that we may not conjoin propositions individually known to be true. Thus if I know that p and I know that q, I do not, according to Goldman, thereby know that p & q. Goldman rejects the 'conjunctivity of knowledge.' So if I know that ticket 1 will not win, . . . , and I know that ticket $n+1$ will not win, it does not follow that I know that no ticket will win. In the presentation of his position Goldman acknowledges several assumptions. Foremost among them is that the justified true belief account of knowledge is correct. He also endorses the more widely held claim, which he refers to as 'principle (A),' that one can be justified in believing a false proposition.

On the supposition that knowledge is justified true belief, there is only one ingredient in knowledge other than justified belief. That ingredient is truth. Conjunctivity holds universally for true propositions. So if conjunctivity fails for knowledge, it does so because it fails for justified belief. Therefore on the assumption that knowledge is justified true belief, the question of the conjunctivity of knowledge reduces to that of the conjunctivity of justified belief. Goldman seeks to answer these questions in the negative. He begins with an argument on pp. 7–8, which we shall call 'argument 1.' He states that a set of falsely believed propositions {P1, P2} may have justification equal to that of a set of propositions {P3, P4} that are known. He infers that if justified belief in P1 and in P2 does not entail justified belief in the conjunction P1 & P2, then knowing P3 and knowing P4 does not entail justified belief in the conjunction P3 & P4, 'since the justification and rationality of belief in both sets is by hypothesis the same' (p. 8).

---

[1] 'A Note on the Conjunctivity of Knowledge', ANALYSIS 36.1 (1975), pp. 5–9. The following restatement of Goldman's argument makes implicit use of his connection between rational and justified belief (p. 8, n. 1).

Argument 1 begins with the statement, 'But the retention of (A) is incompatible with the Conjunctivity of Knowledge.' As we have just seen, Goldman then uses his assumptions to establish the conditional proposition

> (C) if conjunctivity fails for justified belief, then it fails for knowledge.

But the statement just cited implies that Goldman intended to prove the consequent of (C), not merely the conditional. In the next paragraph Goldman begins another argument, drawn from a pair of examples. Let us call it 'argument 2.' He introduces it as follows. 'The incompatibility of (A) and the conjunctivity of knowledge can be seen less abstractly by example' (p. 8). Beyond any question, Goldman believes that the more abstract argument 1 has already demonstrated this incompatibility. But in fact, no such thing has been shown. Neither the antecedent nor the consequent of (C) is supported in argument 1. Why may we not infer that the conjunction P1 & P2 is justified? It is false. But its falsity cannot preclude the inference, since Goldman's endorsement of (A) explicitly commits him to the possibility of justified false beliefs. Goldman provides, in argument 1, not a shred of evidence in support of either clause of (C).

Argument 2 presents a pair of interesting and suggestive examples. It is intended to illustrate concretely what Goldman believes has already been proven, and to subvert the intuitive support that has been claimed for the conjunctivity of knowledge. It is put forth as an independent argument against the conjunctivity of knowledge. Both examples conclude with a denial of that principle. However, the most they can be construed as proving is the antecedent of (C). Thus argument 2 is dependent upon argument 1, which establishes (C). But much more importantly, we must note that the connection between the cases described and the antecedent of (C) cannot be completed without the aid of several further premises, to which Goldman makes no reference. For example, a number of assumptions must be made concerning the existence and nature of a probabilistic rule of acceptance. Argument 2 requires also the controversial claim that beliefs either are, or behave like, subjective probabilities.

The examples in argument 2 make reference to a 'reliable method' that produces accurate results with a frequency of x, where $\cdot 95 < x < 1$. The examples also invoke a rule that requires a probability of y for acceptance. Let us say $\cdot 90 < y < \cdot 95$. Any result reached by the reliable method will meet the acceptance criterion, and thus may be justifiably believed. If we conjoin n results, the degree of reliability of the conjunction will be x^n. But when n is large enough, x^n will be below the acceptance level. We are then not justified in believing such conjunctions. So it appears that conjunctivity fails for justified belief. However, the facile conflation of beliefs and subjective probabilities, which is presupposed in this argument, has been successfully challenged by Isaac Levi (*Gambling with Truth,* Cambridge, Mass.: M.I.T., 1973, pp. 122–24). It has been shown by Levi that there is at least one standard sense of 'belief' according to which beliefs do not quantitatively combine as prescribed by the probability calculus. Rather, the degree to which one believes a conjunction is equal to the least degree to which one believes any of its conjuncts. Hence, the situation described in argument 2, in which the degree of belief in a conjunction dwindles below that of any conjunct, cannot arise.

Goldman concludes by pointing out that the 'conjunctivity of rational belief appears far less plausible to the degree that its stronger counterpart fails' (p. 9). It is true that the converse of (C) follows directly from the assumption that knowledge is justified true belief. But as we have seen, Goldman argues against the conjunctivity of knowledge only by arguing against the conjunctivity of rational belief. To imply that the rejection of the conjunctivity of knowledge, thus derived, casts new doubt on the conjunctivity of rational belief is indeed to argue in a circle. In any case, the arguments presented by Goldman defeat neither the conjunctivity of knowledge nor of justified belief. Argument 2 leaves untouched a number of essential steps. Without an acknowledgement and defence of these additional premises the examples are open to various alternative interpretations. And argument 1, although it establishes (C), is construed by Goldman, quite unaccountably, as a proof that conjunctivity fails for knowledge. Thus the principle of conjunction retains its plausibility in epistemic contexts.

The Johns Hopkins University

© Fred Chernoff 1980

CARTESIAN MADNESS

By MICHAEL WELBOURNE

IN the recent discussion of Descartes's reference to madness in the First Meditation (Steven DeHaven, 'Descartes, Frankfurt and Madness', ANALYSIS 38.2 March 1978; Robert Fahrnkopf, 'Cartesian Insanity', ANALYSIS 39.2 March 1979) I believe that the crucial point has been lost through mistranslation.

I

What is it that Descartes says he would have to be mad to doubt? He writes: 'Manus vero has ipsas, totumque hoc corpus meum esse, qua ratione posset negari?' (AT vii 18). Anscombe and Geach translate as follows: 'Again, these hands and my whole body—how can their existence be denied?' (AG 62). They are followed by DeHaven and, tacitly, by Fahrnkopf. Their version, incidentally, is endorsed by quotation in Kenny (*p.* 15) who nevertheless gives the reference to Haldane and Ross (HR i 145).

This is a strained reading of the Latin; it is much more natural to take the 'meum' as the complement of 'esse', and this more natural reading gets support from the fact that the thought expressed occurs in a context where doubts about *existence* are not at issue. Our senses, it has just been suggested, may mislead us about very small or rather distant objects but there are many things we cannot sensibly doubt at all. For example . . . Our sentence gives, so to speak, the crowning example. The thought that I might mistakenly attribute existence to some small or distant objects is incoherent and Descartes does not entertain it. I can, however, envisage being mistaken about the properties of such objects —the shape, say, of a distant tower—and it is, surely, this familiar sort of error that Descartes has in mind.

This, then, is the context and what Descartes is saying is: 'How could I sensibly doubt that I am in this place . . . holding this paper in my hands, indeed (our sentence) how could I doubt that these hands and this whole body *are mine*?' The suggestion is that he has some beliefs *about* himself or his body (not yet distinguished) which resist being called into question; pre-eminent among these is the belief that this body is *his*.

The French version, followed, *pace* Kenny, by Haldane and Ross, supports me, but I will not claim credit for it on the grounds that Descartes himself amended and approved it (*cf.* Fahrnkopf *p.* 68) since it is guilty of the opposite mistake just a few lines later. Descartes says: 'Let us suppose that we are dreaming . . . nec forte etiam nos habere tales manus, nec tale totum corpus' (AT vii 19)—'that perhaps even we *have* no such hands, no such body [*sc.* as we dream we have]'. But the

French version takes him to be advancing a doubt about the *properties* of our hands and bodies. ('Et pensons que peutetre nos mains, ny tout notre corps, ne sont pas tels que nous les voyons' AT ix 15; *cf.* HR i 146) On my view this error is hardly less destructive of Descartes's strategy than the first.

II

The matter is important because Descartes *never* deviates from the position that it would be madness for me to doubt *concerning* these hands and this body (= mine) that they are mine. It is not that when he reflects on the possibility of dreaming he sees a way of doubting what formerly he thought could only be doubted on the, as he thinks, mad supposition that he might be mad. There just is no *sane* way for me to doubt concerning my body that it is mine, although the reflection that I may be dreaming can give me a reason to doubt whether I *have* a body and hence whether there is a body to be mine. The peculiarity of this body (for me) is that what teaches me that it exists also teaches me that it is mine. There is no further lesson to learn as I might have to learn, for example, who owns the umbrella which I can see in the corner. *My* body *is* the body I apprehend as mine.

In a similar way when I have a consciousness as of my arm hurting I cannot doubt that my arm hurts save by doubting that I have an arm to hurt (*cf.* the discussion of the phantom limb—AT vii 76–7; HR i 189; AG 113). The parallel is unsurprising since it is precisely this sort of consciousness which leads me to believe both that my arm *hurts* and that what hurts is *my* arm. These matters are of central concern in the Sixth Meditation (*cf.*, for example, the famous passage about the pilot in the ship—AT vii 81; HR i 192; AG 117). It is a further argument in favour of my interpretation that it would be exceptional if that concern were not somehow foreshadowed in the development of the doubt.

Thus I agree with Fahrnkopf that dreaming and insanity are not equivalent in Descartes's eyes and that for better or worse he will not allow the possibility of madness a role in establishing the doubt. For all that, I think that Fahrnkopf has not seen to the heart of the matter. The point is not that there is a certain proposition, 'this body (= mine) exists' which I can sanely doubt only on the supposition that I may be dreaming. Rather it is that the proposition 'this body (= mine) is mine' could *never* be, for me, so long as I am sane, the *direct* focus of a doubt because what makes this body mine is my being conscious of it *as* mine. There would, then, be a peculiar incoherence in my denying that it is mine. It is this incoherence which, in Descartes's eyes, warrants the comparison with the mad.

University of Bristol

© MICHAEL WELBOURNE 1980

REFERENCES

[1] AT: Charles Adam and Faul Tannery, editors, *Oeuvres de Descartes* (Paris 1913).
[2] HR: Elizabeth Haldane and G. R. T. Ross, translators, *The Philosophical Works of Descartes* (Cambridge 193[illegible]).
[3] AG: Elizabeth Anscombe and Peter Thomas Geach, translators, *Descartes: Philosophical Writings* (Edinburgh 1954).
[4] Kenny: Anthony Kenny, *Descartes, A Study of his Philosophy* (New York 1968).

FOR WANT OF A NAIL

By E. J. LOWE

ARE causal relations transitive? On first reflection it might seem indisputable that 'x caused y' and 'y caused z' together entail 'x caused z'. But that this is incorrect is a lesson which may be drawn from the well-known nursery rhyme (*Oxford Dictionary of Nursery Rhymes*, edd. Iona and Peter Opie, Oxford 1951, p. 324):

> For want of a nail the shoe was lost,
> For want of a shoe the horse was lost,
> For want of a horse the rider was lost,
> For want of a rider the battle was lost,
> For want of a battle the kingdom was lost,
> And all for the want of a horseshoe nail.

Each line of the rhyme up until the last quite plausibly asserts that the loss of one thing caused the loss of another; the paradoxical concluding line asserts that the mere loss of a horseshoe nail caused the loss of an entire kingdom, which is most unplausible. (If any of the intermediate lines seems insufficiently plausible, it can easily be replaced by a number of more plausible lines while still preserving the transition from the loss of the nail to the loss of the kingdom.)

Now if, following Mackie (*The Cement of the Universe*, Oxford 1974, pp. 29ff.), we analyse 'x caused y' in terms of the counterfactual conditional 'y would not have occurred if x had not occurred' then, it seems, the rhyme may be more perspicuously, if more prosaically, paraphrased as follows (Mackie's analysis has some further refinements which I have deemed irrelevant here):

> If the nail had not been lost, the shoe would not have been lost;
> If the shoe had not been lost, the horse would not have been lost;
> If the horse had not been lost, the rider would not have been lost;
> If the rider had not been lost, the battle would not have been lost;
> If the battle had not been lost, the kingdom would not have been lost;

So the kingdom would not have been lost if the nail had not been lost.

Expressed in this form the fallacy should no longer particularly surprise us, since it has for some time been recognized that for counterfactual conditionals (unlike material conditionals and strict conditionals) transitivity fails (see e.g. D. Lewis, *Counterfactuals,* Cambridge 1973, pp. 32 ff., and R. C. Stalnaker, 'A Theory of Conditionals', in *Studies in Logical Theory,* ed. N. Rescher, Blackwell 1968). But what is the source of the fallacy? Why does transitivity fail for causal sentences? Lewis contends that, for 'would' counterfactuals quite generally, where the inference from $\chi \mathrel{\Box\!\!\rightarrow} \phi$ and $\phi \mathrel{\Box\!\!\rightarrow} \psi$ to $\chi \mathrel{\Box\!\!\rightarrow} \psi$ fails, 'the antecedent of the first premise must be more far-fetched than the antecedent of the second', and adds that 'in . . . all . . . counterexamples to transitivity, the "might" counterfactual $\phi \mathrel{\Diamond\!\!\rightarrow} \sim\chi$ is true' (Lewis p. 33).

Does this explain the non-transitivity of causation illustrated by the nursery rhyme? It is not clear to me that it does. We may give the inference expressed in our paraphrase of the rhyme the following symbolic representation:

$$\begin{array}{ll} \text{A} & p \mathrel{\Box\!\!\rightarrow} q \\ & q \mathrel{\Box\!\!\rightarrow} r \\ & r \mathrel{\Box\!\!\rightarrow} s \\ & s \mathrel{\Box\!\!\rightarrow} t \\ & t \mathrel{\Box\!\!\rightarrow} u \\ \cline{2-2} & p \mathrel{\Box\!\!\rightarrow} u \end{array}$$

This inference may presumably be broken down into the following series of subsidiary inferences:

$$\begin{array}{llllllll} \text{B} & p \mathrel{\Box\!\!\rightarrow} q & \text{C} & p \mathrel{\Box\!\!\rightarrow} r & \text{D} & p \mathrel{\Box\!\!\rightarrow} s & \text{E} & p \mathrel{\Box\!\!\rightarrow} t \\ & q \mathrel{\Box\!\!\rightarrow} r & & r \mathrel{\Box\!\!\rightarrow} s & & s \mathrel{\Box\!\!\rightarrow} t & & t \mathrel{\Box\!\!\rightarrow} u \\ \cline{2-2} \cline{4-4} \cline{6-6} \cline{8-8} & p \mathrel{\Box\!\!\rightarrow} r & & p \mathrel{\Box\!\!\rightarrow} s & & p \mathrel{\Box\!\!\rightarrow} t & & p \mathrel{\Box\!\!\rightarrow} u \end{array}$$

Now, if inference A fails, it must presumably be because one or more of the subsidiary inferences B—E fail (though there is no reason to suspect any one of these rather than any other). However, according to Lewis, if any one of these subsidiary inferences fails, then 'the antecedent of the first premise must be more far-fetched than the antecedent of the second'. But it is not at all evident that the counterfactual retention (non-loss) of the nail is more 'far-fetched' (or, as Lewis glosses this, 'farther from actuality') than the counterfactual retention of the shoe, the horse, the rider or the battle. Indeed, if anything the counterfactual retention of these items becomes increasingly 'far-fetched' as we proceed along the series: given that each of the items was in fact lost, we are taken increasingly 'farther from actuality' in supposing that they were not. The whole charm of the rhyme resides, after all, in its suggestion that a catastrophe arose from a trivial event.

Lewis' explanation also requires the truth of one or more of the 'might' counterfactuals 'If the shoe had not been lost, the nail might have been lost', 'If the horse had not been lost, the nail might have been lost', etc. (or, what amounts to the same thing, the falsity of one or more of the 'would' counterfactuals 'If the shoe had not been lost, the nail would not have been lost', 'If the horse had not been lost, the nail would not have been lost', etc.). It is not particularly obvious that this condition is satisfied. Indeed, our intuitions provide, it seems to me, no clear guidance here, so that the applicability of this feature of Lewis' explanation cannot easily be established. But I must confess that I have no ready alternative explanation to offer. (In his 1973 paper 'Causation', reprinted in *Causation and Conditionals,* ed. E. Sosa, Oxford 1975, Lewis defines 'causal dependence' in such a way that 'if c and e are actual events, then . . . e depends causally on c iff, if c had not been, e never had existed' (p. 187: a deliberate echo of Hume). This is equivalent to our Mackiean analysis of 'x caused y'. Lewis explicitly allows that causal dependence is non-transitive, but defines *causation* in terms of chains of causally dependent events in such a fashion as to secure its transitivity. At the end of the paper Lewis argues that causal dependence may fail of transitivity on account of 'pre-emption', a type of causal overdetermination: but it is not certain that this could be appealed to to explain why, in Lewis' terms, the loss of the kingdom is not causally dependent on the loss of the nail, nor is it clear how this explanation of non-transitivity would fit in with the explanation given in *Counterfactuals.*)

© E. J. Lowe 1980

THE TRANSITIVITY OF COUNTERFACTUALS AND CAUSATION

By J. L. MACKIE

E. J. Lowe's interesting paper ('For Want of a Nail', ANALYSIS 40.1, 1980) shows how people's intuitions may vary. I find the conclusion that the loss of a horseshoe nail caused the loss of a kingdom quite acceptable, and equally the counterfactual paraphrase of this, that the kingdom would not have been lost if the nail had not. This conclusion is surprising if one assumes that causes must be proportional to their effects, but the moral of the nursery rhyme is that small causes can trigger very large effects. Why not? Of course, these will be only partial causes, there must be other factors involved, other forces released by these triggers. But that is normal: most of what we recognize as causes are only partial causes. So I find in this example no proof of the non-transitivity of either causation or counterfactuals.

It is true that Lewis, Stalnaker, and others have argued that transitivity fails with counterfactuals. This is a consequence of their analysing a counterfactual 'If it had been that *p*, it would have been that *q*' as (roughly) 'In the nearest possible world where *p*, there *q* also'. For while *r* may hold in the nearest possible world where *q* holds, and *q* in the nearest one where *p* holds, *r* may fail to hold in the nearest one where *p* holds. The nearest world where *p* holds may not be as near as some other one where *q* holds, and *q* may bring *r* with it in the latter but not in the former.

These theorists are, in effect, reading each counterfactual with some such implicit qualification to its antecedent as 'in the easiest way'. Clearly, from 'If *p* in the easiest way, then *q*' and 'If *q* in the easiest way, then *r*' we cannot validly derive 'If *p* in the easiest way, then *r*'. Thus explained, the non-transitivity of counterfactuals is nothing surprising: it is merely a variant of a very old form of fallacy, a syllogism with an ambiguous middle term. If '*q*' as the antecedent of the second premiss is short for '*q* in the easiest way', while '*q*' as the consequent of the first premiss has no such qualification, there is no valid hypothetical syllogism. On the other hand, if '*q*' in the second premiss is read as covering the way of its being *q* that is indicated by the consequent of the first premiss, then there is a valid hypothetical syllogism; transitivity holds. All the so-called 'counterfactual fallacies', all the alleged differences in logic between counterfactuals and other conditionals set out, for example, by Lewis on pp. 31–6 of *Counterfactuals,* have a similar origin.

We could introduce such a failure of transitivity into the horseshoe nail example, though at the price of some artificiality. Suppose that at the historical period in question semi-miraculous events were common.

Then perhaps the nearest possible world in which the rider was not lost is one in which after his horse became lame he was carried into the battle in a whirlwind; this hint of supernatural support would have so encouraged the troops that they would have won the battle, whereas the mere arrival of one extra horseman in the ordinary way would not have made this difference. So 'If the rider had not been lost', read in Lewis's style, will here mean 'If this rider had arrived by whirlwind', whereas 'the rider would not have been lost' in the third premiss will mean only 'the rider would have turned up normally on a sound horse'. With these readings the premisses are all true, but not the conclusion. It is of such a case that Lewis says 'the "might" counterfactual $\phi \Diamond\!\!\rightarrow \sim\chi$ is true'—here 'If the rider had not been lost the horse might have been lost', which is correct.

Of course none of this applies to the story intended in the nursery rhyme. As Lowe says, Lewis's explanation of transitivity failure does not work here. But is there any failure to be explained?

Well, there may be. Let us go back to causation. Part of the meaning of a singular causal statement seems to be that the cause was necessary and sufficient in the circumstances for the effect. Now if p is necessary and sufficient for q in some circumstances, and q is necessary and sufficient for r in other circumstances, it follows that p is necessary and sufficient for r in the circumstances only if these last circumstances are taken to include both the former sets. The conclusion of the horseshoe story can be read as referring to a total set of circumstances including all those in relation to which the various premisses hold in turn: then causation is here transitive, and the kingdom *was* lost for want of a horseshoe nail. But there are two ways in which transitivity might be thought to fail here. First, though each premiss indicates well enough the sort of circumstances in relation to which it holds, the conclusion, taken on its own without the premisses, gives no such indication, and might be initially unacceptable for that reason, though it would become acceptable when the premisses were stated and the full set of circumstances thereby indicated. Secondly, though a singular causal statement ordinarily specifies only a partial cause, it might be considered acceptable only if it picks out a cause which can be seen as a significant portion of the total cause: perhaps, if we could measure this, it would need to be at least one-tenth of the total cause. If so, causation is not transitive. The cause-factor mentioned in each of the five premisses could be at least one-tenth of the total cause at that step, and yet the loss of a nail might be little more than one hundred-thousandth of the total cause of the loss of the kingdom.

University College, Oxford

© J. L. Mackie 1980

ARISTOTLE ON THE SEA-BATTLE: A CLARIFICATION

By Malcolm F. Lowe

I

IN the controversy of recent decades over Aristotle's intentions in the ninth chapter of his *De Interpretatione,* a certain amount of common ground has been established.[1] Thus it has been recognised that Aristotle concedes to his philosophical opponent(s) the necessary truth of more statements than might a non-determinist today, and makes a stand only on the contingency of statements about particulars in the future (such as whether a sea-battle will take place tomorrow, or whether his cloak will be cut up before being worn out).

It is also recognised that Aristotle considers two distinct arguments advanced by those who hold even these statements to be necessarily true or necessarily false. The first argument (18a 34-b 9) bases itself on the claim that if two people make *contradictory* statements about such a future particularity, one or other must be telling the truth. The second (18b 9–16) claims that if it *was always* true to say that what eventually happens would happen, then one could not have said so truthfully without it happening.

What does not seem to have been noticed is that Aristotle, in responding to his opponents (19a 23-b 4), answers these two arguments *separately*.[2] The obscurity is due to a peculiarity of the passage where Aristotle states his basic answers (19a 23–29). This passage consists of two sentences linked by the phrase 'the same account [holds]' (*ho autos logos*), although it is not immediately evident that he *is* talking about the same thing in both sentences. Let us, therefore, simply leave out the phrase in question[3] and consider the two sentences separately:

(a) 'Necessarily, that which is is when it is, and that which is not is not when it is not; but neither is everything necessarily which is, nor is not that which is not; for it is not the same [to say] that everything which is is of necessity *when* it is, and that it *without qualification* is of necessity; and likewise concerning that which is not.' (19a 25–7)

---

[1] Of the extensive literature on this topic, the two items most relevant to what follows are: J. L. Ackrill's translation and commentary, *Aristotle's Categories and De Interpretatione* (1963); and J. Hintikka, 'The Once and Future Sea Fight' (1964), slightly revised (with further bibliography) in his *Time and Necessity* (1973). Both accounts seem to me to be substantially correct, except inasmuch as they lack precisely the clarification which will be made here. The translations given here are indebted to those of Ackrill, though not identical with them.

[2] For instance, recently C. J. F. Williams claimed, in 'True Tomorrow, Never True Today', *Philosophical Quarterly* 28 (1978), 285–299, that the second argument is nowhere answered by Aristotle (pp. 286 ff.). On the other hand, Hintikka (in *op. cit.*) took Aristotle to be answering *above all* the second argument, though recognizing that there are two distinct arguments (*ibid.*, p. 168).

[3] The phrase is actually missing from codex Ambrosianus L93 (one of the two chief mss. underlying Minio-Paluello's edition of the *Categories* and the *De Interpretatione*).

(b) 'And concerning the *contradiction*: it is necessary for everything to be or not to be, and that it indeed will be or will not; but one cannot divide and say that the one is necessary' (*dihelonta ge eipein thāteron*—i.e., *to heteron—anankaion*, 19a 27–9).

As soon as one is no longer required to regard these two sentences as making the same (or a similar) statement, it is possible to see that the *first* of them is Aristotle's basic answer to the *second* argument of his opponents, while the *second* sentence answers their *first* argument. Such a relationship can be seen, to begin with, simply in their verbal form: it is sentence (b) and the first argument which are concerned with a pair of *contradictory* statements about an event, while it is sentence (a) and the second argument which turn upon the time *when* the occurrence of the event is discussed.

The relationship is confirmed when one examines the arguments in detail and notices at what point Aristotle intends to block further moves. The first argument starts with a statement of what is to be proved (18a 34–35): 'For if every affirmation or denial is true or false, it is also necessary for everything either to be the case or not to be the case.' Now the argument proper begins: 'For if one person says that something will be and another says that the same thing will not, clearly it is necessary for the one of them to be telling the truth (*anankē alētheuein ton heteron autōn*) . . . ' It is at this point that Aristotle interrupts with the cry 'Enough!'—since the principle expressed by sentence (b) is about to be violated (note the verbal parallel with sentence (b)). For, he says (19a 30–2), 'it is necessary, for example, that there will be a sea-battle tomorrow or there will not be, but', since dividing is forbidden by sentence (b), 'it is not necessary for a sea-battle to happen tomorrow nor [for one] not to happen, though to happen or not to happen is necessary.' And, consequently (19a 33 ff.), 'since it is likewise with the truth of statements as with [the way] things [are]', you also cannot assert that, if one person affirms and another denies that a sea-battle will happen, either the affirmation is necessarily true or the denial is necessarily true.

Similarly with the second argument. This begins (18b 9 ff.): 'If [something] is white now, it was true to say earlier that it will be white, so that it was always true to say of anything that happens[1] that it will be; but if it was always true to say that it is or will be, this also cannot not be or not be going to be . . . ' Here again Aristotle cries 'Enough!'—this time in terms of sentence (a): everything necessarily is *when* it is (so that what is white now is now necessarily white), but you cannot jump to asserting its necessity even without the qualification 'when it is', and (specifically) to its having been already necessary when someone forecast it in the past. Similarly, one may now make a true forecast about the future, but this does not make what is already truly forecast already necessary.

---

[1] Reading with Ammonius and the same codex Ambrosianus L93; other ancient authorities have 'has happened'.

II

We have seen, therefore, that Aristotle is blocking two separate arguments by making two separate distinctions. So in what sense, if any, do the two distinctions—sentences (a) and (b)—represent 'the same account' (to return to the problem which we mentioned at the beginning)? The commentator Ammonius believed that the link lies in the distinction between two kinds of necessity which is mentioned in sentence (a): between the necessity of something's being when it is and necessity without qualification. Aristotle's opponents, namely, argue for the necessity of all events by starting either from the first kind of necessity in application to present events (their second argument), or from the second kind of necessity as manifested in the Law of Excluded Middle (their first argument).[2]

Another possible explanation of the phrase 'the same account' is suggested by something that Aristotle says at the end of the passage in which he answers his opponents. This passage (19a 23–b4) divides as follows: first come sentences (a) and (b), then the detailed application of (b) to their first argument (the application of (a) to their second argument is not spelt out), and lastly the remark: 'For it is not likewise concerning things that are and concerning things that are not [i.e., at the moment] but are capable of being or not being, but as has been said.' If this is a reference back to sentence (a), what it suggests—in its context—is that sentence (b) is related to sentence (a) in the sense that (b) states a prohibition which can be effective *only* in respect of non-present events. For since Aristotle has conceded in sentence (a) that what is is necessarily when it is, it follows that if one person affirms and another denies that a sea-battle is taking place *right now*, then indeed either the affirmation is necessarily true at the moment or the denial is, so that in *this* case one can ignore sentence (b) with impunity.

On either of these explanations, Aristotle is shown to be not so much 'confused' (thus Ackrill, *op. cit.,* p. 138) as rather 'confusing'. If we follow Ammonius, he is saying that two basically unrelated distinctions constitute 'the same account' because both are concerned with false attributions of necessity. Alternatively, he means that they are unrelated *except* inasmuch as the first distinction presupposes a restriction upon the force of the second. In neither case, however, are the two distinctions shown to be the same or to be derivable from a common underlying distinction.

---

[2] Ammonius, *In Lib. De Int.* (ed. Busse), 153–154. He seems, however, to have regarded the two arguments as variants of each other (*ibid.,* 139–145). In all this the approach of Hintikka (*op. cit.*) is very similar.

III

While the foregoing analysis is deliberately modest in scope, it does answer one of the questions constantly raised about Aristotle's intentions: he is not denying the application of the Law of Excluded Middle to statements about future events of some kind, but only the inference from that law to the necessity of all events. In other words, I argue for what N. Rescher has called the 'medieval' as opposed to the 'orthodox' interpretation of Aristotle's doctrine. Rescher's terminology, however, is mistaken and should be abandoned, since he failed to notice that Ammonius and Boethius very clearly adopted the so-called 'medieval' interpretation and are almost certain to have been its sources. For the same reason, one cannot follow Hintikka (*op. cit.*) in terming the denial of the Law of Excluded Middle the 'traditional' interpretation.

'As the medievals standardly put it', according to Rescher, future contingents 'are not *determinately* true or *determinately* false'; this corresponds to the assertion of Ammonius that of a pair of contradictory statements of this kind 'one cannot yet say determinately [*hōrismenōs*] which will be true and which will be false'.[3] Boethius too made it emphatically clear that the Law of Excluded Middle had to be retained; indeed, this line of interpretation probably goes back at least as far as Porphyry, whom Boethius explicitly claimed to be following for most of his exegesis of *De Interpretatione* (and on whom Ammonius chiefly relied as well).[4] It was as standard an interpretation among Aristotle's commentators in late antiquity as among his medieval ones, and thus as 'orthodox' and 'traditional' as one could desire. The contrary interpretation is also ancient—but has its source rather among Aristotle's opponents. For instance, Boethius (ed. Meiser), II, 208, 1–18 criticises the Stoics and some others, who 'thought Aristotle to say that future contingents are neither true nor false'.

---

[3] Rescher, *Essays in Philosophical Analysis* (1969), p. 275; Ammonius, *op cit.*, 130, 23–24. See also *ibid.*, 130–155, *passim* (for repeated uses of *hōrismenōs* and *aphōrismenōs* in this sense), and esp. 139, 13–20, where Ammonius states that of the contradictory pair one will be true and one false, *ou mentoi aphōrismenōs all' aoristōs*. In alleging that Ammonius and Boethius rejected the Law of Excluded Middle, Rescher (*op. cit.*, p. 274, n. 5) gives incomplete quotations which omit precisely the words *aphōrismenōs* and *definite* respectively.

[4] On the sea-battle, for instance, see Boethius (ed. Meiser), I, 123, 5–10: '. . . it is indeed necessary [for it] to be or not to be, but it is not necessary [for it] to be nor again is it necessary [for it] not to be, but only either to be or not to be. Therefore the whole contradiction will indeed have one part true, the other false, but of them the one will not be definitely [*definite*] true, the other definitely false.' For his general indebtedness to Porphyry, see *ibid.*, II, 7, 5–7 (and confirmatory remarks in the sequel). On Ammonius' sources and subsequent universal influence, see Busse's introduction to his edition.

Van Leer Foundation, Jerusalem

© MALCOLM F. LOWE 1980

THE SELF FROM TIME TO TIME

By D. H. MELLOR

FROM the fact that a token of the sentence type

(1) It's two o'clock now

can be used to give valuable information, Zemach (ANALYSIS 39.3, June 1979) concludes that selves have only momentary existence. While I share Bernard Williams' (1973, ch. 6) distaste for immortality, I confess my present self still hopes for more than the mayfly's span of life, and Zemach does not persuade me I can't have it. Believing moreover that others, perhaps even some future Zemach-selves, may wish to share my hope, I hasten now to reassure them. A less bizarre conclusion can and should be drawn from the familiar facts and arguments to which Zemach redirects our attention.

The truth conditions of any token of (1) are not in dispute. A token of it is true if it is produced at two o'clock; otherwise it is false. The time of utterance completely fixes its truth value. There is no scope for anything else to affect the truth value of tokens of (1). Specifically, the pastness, presentness or futurity of the time, two o'clock, to which they refer cannot affect their truth value, even though (1) appears to say that two o'clock has the property of presentness. And the same goes for tensed statements of any complexity, as the stock "semantics" for tense logic (e.g. McArthur 1976) makes clear, albeit unintentionally. This semantics shows, if nothing else, how the temporal truth conditions of a tensed statement are fixed entirely by how much earlier or later its utterance is than various other events and times explicitly or tacitly referred to in it. The supposed properties of pastness, presentness or futurity, whether of these events or times (or, as tense logicians prefer, of the facts of their occurrence) have no scope at all to enter into the truth conditions of any tensed utterance. And since their doing so would provide the only reason for supposing any time, event or fact really to possess these properties, we see that there is no such reason.

Since, moreover, McTaggart (1908) anticipated Zemach in showing the contradictions inherent in this supposition, we have had the best of reasons to deny it. Nothing can be distinguished in reality as being past, present or future—except in the tenseless relational sense in which events later than a time may be said to be "future" to it, earlier events "past" at it, and simultaneous events "present" at it. Nor therefore can any other real nonrelational distinction be consistently supposed to rest on these illusory ones. In particular, as Zemach observes, 'it is nonsense to say that "the past" or "the future" are not fully real.' It is also nonsense to distinguish the past from the future in other ways still current among philosophers: as being objectively fixed (as opposed to open), actual (as opposed to possible), unique (as opposed to multiple), or certain (as opposed to chancy). All this is nonsense, and it's time it stopped.

Tense, in short, is not an aspect of reality, but the way the truth conditions of tokens of certain sentence types depend on when they are produced; just as spatial token-reflexive expressions ('here', 'sixteen miles North East') make the truth conditions of tokens of other sentence types depend on where they are produced. Now to these sentence types there correspond definite states of belief (that it is now two o'clock, that Ely is sixteen miles away North East) whose truth conditions likewise depend on when or where they are held. The truth conditions of tenseless beliefs (e.g. that the 1979 Cup Final starts at two o'clock) and their spatial equivalents (e.g. that Ely is sixteen miles North East of Cambridge) do not, by contrast, depend in this way on when or where they are held. A tensed belief state is therefore never identical with a tenseless one, and consequently neither are the meanings of the sentences expressing them. There is, as Prior and others have rightly maintained, no tenseless *translation* of tensed discourse, even though there is no tense in the reality that determines whether it is true. There is no special temporal mystery about this; it merely illustrates the fact that sentences with token-reflexives cannot be properly translated by sentences without them, whether the token-reflexives be temporal, spatial or personal.

All this, *pace* Zemach, has nothing to do with selves, except insofar as sentient selves are needed to supply the conventions which make linguistic types out of specific types of sound and of inscription. Tokens of (1) need not be uttered by a self; most of them are in fact given out by clocks and not noted by anybody. They could quite well continue to be given out after all our selves are dead. The things books and road signs say would surely retain their truth value in those circumstances, and if they would, so would the deliverances of Big Ben. Now no one, I trust, thinks Big Ben is a self, let alone a succession of momentary selves; it is rather too far down the Chain of Being. So it couldn't truly say Zemach's

(5) I am located at two o'clock.

But it can, and repeatedly does, truly say (1); which therefore does not, as Zemach asserts, have (5) for its content.

The fact is that Zemach has misplaced the connection between tense and the self. It really lies in the fact that agents need beliefs whose truth conditions change with time in the way those of tensed beliefs do. Since it is the object of all belief to be true, we have to change our tensed beliefs from time to time in order to try and keep them true, and for this purpose we need what is in effect an internal clock. Thus if my tensed belief about two o'clock is to remain true, it must change at two o'clock to (1) from the belief state I was previously in, namely that two o'clock is still future. And this is the very change that is needed to cause me to do at two o'clock whatever it is I desire to do then. As Perry (1979) has shown, nothing else will do this. Certainly no amount of unchanging tenseless belief, however relevant and true and however combined with specific desire, will prompt me on its own to timely action.

Tensed beliefs are thus indispensable for causing timely action, and therein lies the value of true tokens of (1). I may want to do something at two o'clock, and to do it I need to believe (1) at that time. But this belief does not automatically arrive then; our internal clocks are far from infallible. The belief may need to be induced, and that is the function of a token of (1). If, when I hear it, it is true, I hear it at two o'clock. So if my hearing it causes me at once to believe what I hear, I thereby acquire the true belief I need to prompt me to do what I want at the right time. The value of a true token of (1) depends therefore on what, if anything, its hearers want to do at two o'clock and whether, but for hearing it, they would have come to believe (1) in time to do it. But that it can be of the greatest value, as Zemach rightly says, is obvious.

This explanation of how tokens of (1) may be valuable does not have to appeal to selves being merely momentary, and that I take to be a virtue of it. Perhaps selves are momentary (though I don't believe that even for a moment), but our use of temporal token-reflexives is not a good reason for thinking so. Where Zemach went wrong was in his rejection of his

(4) The time at which this very token of 'now' is uttered = two o'clock.

He may be right to say that (4) 'is quite unhelpful as an analysis of (1)'; but it does in fact have the same truth conditions as (1). Indeed not only (4), *any* assertion (4*) got from (4) by replacing 'now' with any expression whatever, has the same truth conditions, as the Editor has pointed out to me. The reason is that the work of making the assertion token-reflexive with respect to time (rather than speaker or place), which in (1) is done by 'now', is done in (4*) by 'the time at which this token of': and once that is done, the *type* of the token referred to is immaterial. Whatever it is, tokens of (4*), like those of (1), are true if uttered at two o'clock and false at all other times.

So it is in a sense true, as Zemach says, that '*now* is my time, my only time. I have no other.' But the reason is not that the self is only momentary. It is simply that if, while I am doing (or thinking or experiencing) anything, I simultaneously ask myself when, in tensed terms, I am doing it, the answer has to be 'Now'. This answer-token 'now' is made true just by being simultaneous with whatever the question was about, just as a token of (1) is made true by being simultaneous with two o'clock. That is all the much overrated experience of being confined, in one's perception, thought and action, to the present time amounts to.[1]

Cambridge University © D. H. Mellor 1980

REFERENCES

R. P. McArthur (1976) *Tense Logic*. Dordrecht: Reidel.
J. E. McTaggart (1908) 'The unreality of time.' *Mind* 18, 457–74.
John Perry (1979) 'The problem of the essential indexical.' *Nous* 13, 3–21.
Bernard Williams (1973) *Problems of the Self*. Cambridge University Press.

---

[1] This note was written during my tenure of a Radcliffe Fellowship, for which I am indebted to the Radcliffe Trust.

PROPER NAMES AND DEFINITE DESCRIPTIONS WITH WIDEST POSSIBLE SCOPE

By JAMES HUDSON *and* MICHAEL TYE

ONE popular form of argument against the view that proper names are disguised definite descriptions goes as follows:

(1) If *N* were to mean *the F*, then 'N might not have been an F' would be false.

But

(2) 'N might not have been an F' is true.

Therefore,

(3) *N* does not mean *the F*.

This style of argument, which derives from Kripke's lectures 'Naming and Necessity' (in *The Semantics of Natural Language*, ed. by Davidson and Harman, Dordrecht 1972, pp. 253–355), has been criticized by several defenders of descriptivism (see, e.g., M. Dummett, *Frege: Philosophy of Language*, Duckworth 1973; L. Linsky, *Names and Descriptions*, Chicago 1977; B. Loar 'The Semantics of Singular Terms', *Philosophical Studies* 1976, pp. 372–3). The standard response is to maintain that although premise (2) is true, premise (1) is false, since proper names are to be identified with definite descriptions which take the widest possible scope in the sentences in which they occur.

To see how this response works in a particular case, let us take the name 'Benjamin Franklin' and the description 'the inventor of bifocals'. Here it seems intuitively clear that (2) is true for the given *N* and *F*, since being an inventor of bifocals is evidently a contingent property of Benjamin Franklin. However, on the view that names are descriptions which have widest scope, (1) is false, as may be seen by substituting 'the inventor of bifocals' for 'Benjamin Franklin' in both *N*-positions of (1) and giving the description in the mentioned sentence wider scope than the modal operator. From the fact that 'the inventor of bifocals' means *the inventor of bifocals*, it does *not* follow that the inventor of bifocals is such that he could not have failed to invent bifocals.

It seems to us that the above response is an adequate rejoinder to the stated form of argument. However, we believe that there is a way of revising the premises of that argument which makes the appeal to scope distinctions irrelevant. We turn next to explaining how this is so.

Consider the sentence

The F is an F

(for any syntactically appropriate value of 'F'). This is not itself a necessary truth, but it is entailed by the sentence 'There is exactly one F' (given a fixed value of 'F'). Let us say, then, that it is *necessary on the assumption that there is a unique* F, meaning by this simply that it is a necessary consequence of 'There is exactly one F'. Now, it seems to us, if *N* were to mean *the F*, then *N* would be substitutable *salva veritate* for *the F* in '"The F is an F" is necessary on the assumption that there is a unique F'. Hence we have

(4) If *N* were to mean *the F*, then 'N is an F' would be necessary on the assumption that there is a unique *F*.

Conjoining (4) with

(5) 'N is an F' is not necessary on the assumption that there is a unique *F*,

we may deduce that *N* does not mean *the F*. This move from (4) and (5) to (3) is our proposed way of modifying the argument.

The advantage of our revision over the original is as follows. Premise (4) is much more solidly based than is (1), since the scope ambiguity which infects the latter is absent from the former. This is the case because, while in (1) we have a remark about a sentence which contains a modal operator, in (4) we have rather a modal remark about a non-modal sentence. Furthermore the assumption upon which (4) is based—that synonyms may replace one another without changing the modal properties and relations of the sentences which contain them—is generally accepted: disputes about the substitutivity of synonyms arise not in modal contexts but elsewhere. And premise (5) is really no more contentious than (2). To see this, let us return to our example, specifically to the sentences 'Benjamin Franklin is an inventor of bifocals' and 'There is exactly one inventor of bifocals'. Since someone other than Franklin might have invented bifocals (and done so without the assistance of anyone else), the former sentence is not a necessary consequence of the latter, and so it is not necessary on the assumption that there is a unique inventor of bifocals. Hence (5) is borne out for the given *N* and *F*; and clearly this result can be generalized.

It seems to us that our revised form of argument is decisive: the view that proper names are synonymous with definite descriptions cannot be saved by giving the descriptions widest possible scope.

Northern Illinois University © JAMES HUDSON *and* MICHAEL TYE 1980

from Basil Blackwell

Wittgenstein—Understanding and Meaning

An Analytical Commentary on the Philosophical Investigations (Volume 1)

By G. P. Baker and P. M. S. Hacker

Wittgenstein—Understanding and Meaning is an analytical philosophical commentary on Wittgenstein's *Philosophical Investigations*. This first volume covers 1–184, which roughly coincides with the 1937 draft version of the *Investigations*.

The exegesis classifies, analyses and evaluates Wittgenstein's arguments. It draws on the whole of the unpublished papers, paying special attention to the development of Wittgenstein's thought after 1929. Comparisons with the arguments of Frege, Russell and the *Tractatus* are made in order to clarify the background to, and evaluate the cogency of Wittgenstein's later work. This is followed by seventeen comprehensive essays which cover all the major themes in this part of the *Investigations*.

672 pages, £27.50

Central Texts of Wittgenstein

by Gerd Brand

Translated and with an Introduction by Robert E. Innis

This book provides a digest of central texts taken from the whole range of Wittgenstein's work. The texts are organized around a cross-section of topics which informed the total course of Wittgenstein's philosophical writings. In this way Professor Brand has made an ingenious attempt to construct a systematic framework for sampling Wittgenstein.

216 pages, cloth £10, paper £3.95

Wittgenstein and the Vienna Circle

Conversations recorded by Friedrich Waismann

Edited by Brian McGuinness

Translated by Joachim Schulte and Brian McGuinness

This book includes most of the contents of a number of notebooks kept by Friedrich Waismann between 1929 and 1932. In these notebooks he recorded in shorthand views communicated by Ludwig Wittgenstein to Waismann himself and to Moritz Schlick.

The material has considerable historical interest. Most of it is dated, sometimes showing us what Wittgenstein was thinking at times for which we have no other source. It comes from the same transitional period as the *Philosophische Bemerkungen*.

272 pages, £9.95

NOTES

The Analysis Committee consists of: Chairman, P. T. Geach; Secretary, J. H. Benson; T. R. Baldwin, Margaret A. Boden, A. E. Fisher, Andrew Harrison, R. F. Holland, Hidé Ishiguro, J. Kemp, J. H. McDowell, Bernard Mayo, D. H. Mellor, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

Subscriptions. The subscription to Analysis for institutions is £8.00 (inland), £9.80 (overseas), $24.00 (U.S.A.), $28.00 (Canada); for individuals £5.00 (inland), £6.00 (overseas), $15.00 (U.S.A.), $17.50 (Canada). Each volume comprises four numbers, three of 48 pages and one of 64 pages, appearing within the academic year—in January, March, June and October. Orders should be sent to Basil Blackwell, 108 Cowley Road, Oxford OX4 1JF, or placed with any bookseller.

Contributions. **Articles submitted for publication should be addressed to Mr. Christopher Kirwan, Exeter College, Oxford OX1 3DP.** Contributors are asked to note the following requirements.

Articles should **normally** not exceed 3,000 words in length. **Occasionally,** however, longer contributions can be accepted;

They must be type written in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

It is regretted that owing to increased postal charges it is no longer possible to return typescripts unless the following instructions are followed:

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgement is required, a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope **and international reply coupons** of the requisite value, whether for air or surface mail.

Galley proofs of accepted articles will be sent to authors for correction, together with information about offprints. Typescripts will be retained by the Editor on the assumption that authors have kept their own copies.

The copyright of articles printed in Analysis remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.

ISSN 0003–2638

PRINTED IN GREAT BRITAIN BY BURGESS & SON (ABINGDON) LTD., ABINGDON, OXFORDSHIRE

27 MAY 1983

Vol. 40 No. 2 (New Series No. 186) March 198

ANALYSIS

Edited by
Christopher Kirwan

CONTENTS

Report on Analysis problem no. 18	Jonathan Harrison
Swords' points	Margarita R. Levin
Jocasta's crime: a science-fiction reply	Celine Denruyter
Travelling in time	William Godfrey-Smith
Syntactical treatment of negation	Robert K. Meyer
What do the consistency proofs for non-Euclidian geometries prove?	Geoffrey Hunter
Frege on 'possible content of judgment'	Ronald E. Nusenoff
An extension of a proof of Prior's; or, when thinking makes it so	J. J. MacIntosh
Truth and knowability	J. L. Mackie
'If' and 'even if'	John A. Bark[illegible]
Compatibilism and the burden of proof	Peter van Inwagen
Reply to van Inwagen	Richard Foley
Bernard Williams and the Cartesian circle	A. C. Stubbs
Rights, Groups and Discrimination: A Reply to Ketchum	Robert L. Simon

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

BASIL BLACKWELL . ALFRED STREET . OXFORD

New from Basil Blackwell

Wittgenstein's Lectures 1930-1932

Edited by Sir Desmond Lee and John King

The lectures and discussions recorded in this volume are in the main based on notes taken by Sir Desmond Lee and John King in the period from January 1930 to May 1932. They correspond roughly with the period when Wittgenstein was producing *Philosophical Remarks, Philosophical Grammar* and *The Blue and Brown Books,* and deal with such topics as the nature of a proposition, the relation of proposition to fact, the rules of grammar, infinity, the nature of proof, negation, symbols and thought.
160 pages, £7.95

Wittgenstein's Lectures, Cambridge 1932-1935

Edited by Alice Ambrose, from the notes of Alice Ambrose and Margaret MacDonald

The lectures which Wittgenstein gave in Cambridge between 1932 and 1935 represented a radical departure from the work he had published in the *Tractatus*. Topics which Wittgenstein took up during this period include the meaning of a word, its relation to rules for the word's use, general ideas, understanding as a disposition and as a conscious state, propositions, negation, rules of grammar and their relation to fact, and language games.
256 pages, £9.50

The Nature of Philosophy

John Kekes

The whole of this book is informed by the diagnosis which prompted it. This is that we are in urgent need of a systematic and rationally justified worldview and that a humanistically interpreted philosophy is our best hope for providing it. In addition to this overall concern, there are several themes running through the book. One is that philosophy cannot be done as a speciality in isolation from the historical, political, scientific, social, moral, literary, and other cultural influences upon it. The appreciation of cultural influences, however, requires historical understanding. And this means that the construction of worldviews inevitably involves an anthropocentric element. The importance of these humanistic considerations, however, does not change the fact that worldviews must and can be rationally justified. The recurring emphasis upon the necessity and possibility of rational justification in philosophy is another theme of the book.
240 pages, £9.50

ANALYSIS 40:2 MARCH 1980

REPORT ON ANALYSIS "PROBLEM" NO. 18

By JONATHAN HARRISON

THE problem is printed in ANALYSIS 39.2, March 1979, p. 65. There were eleven entries, which I enjoyed reading. I found something interesting in all of them, though most contained errors sufficiently serious to prevent their being published as they stood.

Mr. M. D. Parker argued that time travel would be impossible if time were 'linear', but possible if it were 'planear'. By this he meant that time travel would be possible if I had a different past for every moment of my life (just as a different parallel line, each representing a past, may be run to every point on a line, thus making a plane.) There is, then, not only my present past, but my past pasts and my future pasts, and my 'travelling' to the past consists not in my reaching a present past, but in my reaching a future one. Though I cannot get to my present past (in which I was not, say, at the Great Exhibition) I can bring into existence a future past in which I was. (I found this idea very hard to swallow.) I thought Mr. Parker confused the notion of travelling to a future past with the notion of travelling to another possible world. And I thought that his believing that perhaps one *could* travel to other possible worlds was the product of a great deal of loose contemporary talk about such entities. If one stuck to saying that it is possible that there should be a world in which I (or, which is safer, a man just like me) was at the great Exhibition—and even this does not capture the thought embodied in the words 'It is possible for me to travel to the Great Exhibition'—instead of saying that there *is* a possible world in which I am (was?)[1] at the Great Exhibition, then one would not delude oneself into thinking that a possible world is a place to which one might conceivably go. The reason why Mr. Parker thought that time-travel would not be possible if time were 'linear'—that in this case, I might arrive at a place which was already occupied by someone else—was also unsatisfactory. I cannot get to a past place if someone else was there at that time, but then, neither can I get to a future place if someone else is going to be there at that time.

Mr. Earl Conee made use of the distinction between 'personal' time and 'external' time to try to defend the possibility of Dee's travelling to the past. What happens to him after his arrival 'there' is earlier in external time, but later in personal time, than the events which happened to him before his departure. I did not think that this distinction could do the work that it was intended to do. For surely, when Dee got into a deep freeze at some point in the past, both his overt actions and his private thoughts occurred at the same time. He was feeling apprehensive at the same time that he got in; he could not have got into the deep freeze in A.D. 1600, but be feeling the cold in A.D. 3,000. The fact that, among other things, a man's watch may be running backwards from, say, 1851 to 1971 does not justify us in saying that 1851 is in any sense later than 1971. His suggestion that 'Dee did in Dum' because one section of the Dee-Dum space time worm can kill another is a better idea—though normally one talks of people killing people, not sections of their life-histories doing that.

---

[1] It is worth noting that the word 'was' here cannot have its usual function of indicating that the event being talked about occurred earlier than the time of speaking. Events in possible worlds, or possible events, cannot occur earlier than, later than, or even at the same time as, my utterance.

Miss Harrison Hagan makes the good point that one does not change either the future or the past. What she thinks one does do is make it the case that the future is going to be what it is going to be (and that the past was what it was?). But she goes astray in thinking that in my story Dee displaces Dum (i.e. causes it never to have been the case that Dum ever was married to Miss Jones.) Dee does not displace Dum; he is Dum. Mr. Palmer made the same mistake by supposing that there were two men, one Dum, and the other Dee masquerading as Dum.

P. Swiggers gives a lively and forceful speech for the defence of Jocasta on a charge of bigamy. I would like to have recommended its publication, but I had to fault him both philosophically and professionally. I faulted him philosophically because he supposed that there was an unending series of deep freeze openings followed by marriages followed by births followed by departures, while in fact each event occurs just once. And I faulted him professionally for deliberately pretending, in order to lend respectability to his client, that she was still living with her husband, when he would have known, if he had paid more attention to his brief, that Dum accompanied Dee on his journey, never at any later date to return.

William Godfrey-Smith offered a number of interesting objections to time-travel.

1. There is nowhere to go: supposing that there is somewhere to go is to make the mistake of spatialising time. But one could argue that as, where ordinary travel is concerned, one aims at getting to a place at a time so, where time-travel is concerned, one is aiming at getting to a place as a past time. I would suspect that it was no more than an interesting quirk of the English language that we speak of getting to a certain place by a certain time, and not to a certain time at a certain place, were it not for the fact that we have some choice about the place but (usually) none about the time.

2. Time-travel implies a block universe, and the distinction between personal and external time is misconceived. (I agreed with the second of these statements, but wanted argument for the first.)

3. Dee, unlike other infants, could not have acquired genes from both parents. (But that one acquires genes from two parents is a law of nature, not of logic, so perhaps Aristotle was right about the part played in the genesis of Dee by Jocasta Jones, if he was not right about the part played by any other woman.)

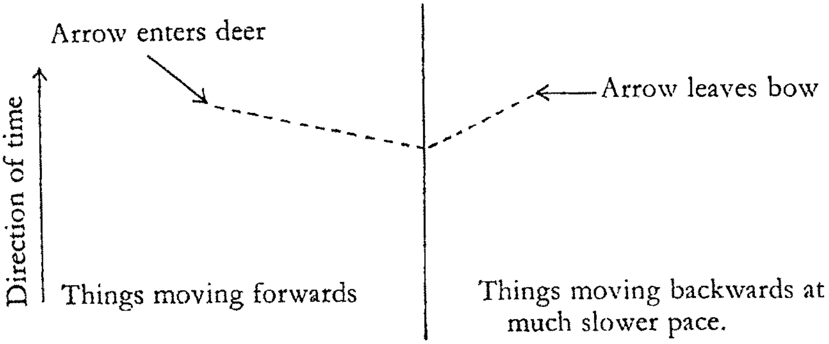
4. If time-travel were to occur, radiant energy would diffuse from the earth to the sun, and the time machine would be invisible. I could not see why he thought either of these things. Why, for example, should the fact that, on the surface of the earth, there is a container holding a clock which is moving backwards and the body of a human being which is getting biologically older rather than younger mean that the container is invisible? For a long time I could not see why a particle moving backwards in time should cause an explosion. For it seemed to me that a particle moving backwards in time from B at *t*1 to A at *t*o, and, say, gaining electrons as it travels, is just the same thing as a particle moving forwards in time from A at *t*o to B at *t*1, and losing electrons. And I thought that the fact that it was possible to describe a forwards moving particle in the backwards moving language could not make it the case that this particle would cause an explosion. This is so. But I understand that there are particles which are such that, if another particle stood in the same relation to them that a man growing biologically younger would stand to a man growing biologically older, there would be an explosion. But it is just a question of empirical fact whether men's bodies contain particles which cannot be 'reversed' in this way—indeed, it is just a matter of empirical fact that they contain particles at all—and my question concerns the logical possibility of time-travel, not its empirical possibility.

5. Dum-Dee's journey 'presupposes the reality of the future in a very strong sense' (strong sense of 'reality'? strong sense of 'future'?) Nor could I see why Mr. Godfrey-Smith should think that Dum-Dee's departure should be any more fixed than anything else about his future. It may be that he has (or they have) to go, because he has (they have) already arrived, but that just means that he goes, not that he does not go deliberately, or that he does not go like a wise man struggling desperately to stay at home. Similarly, if the oracle predicts the murder of Dum's prototype Laius, this does not mean that Oedipus could not help killing Laius. For what the oracle may have foreseen was what actually happened, viz., Oedipus deliberately killing a man who, unknown to him, was his father, and it is not to describe this event correctly to say that Oedipus was unable to do anything other than kill his father.

6. If Jocasta's crime is logically possible, it is logically necessary.

Surely this cannot be. For if it is logically possible for Jocasta to commit incest with one man, it would be logically possible for her to commit it with two; but it could not be both necessary for her to commit it with one and necessary for her to commit it with two. Interesting though Mr. Godfrey-Smith's ideas were, I thought he lacked in care and thoroughness what he had in fertility.

Mr. Godfrey-Smith wonders what would happen on the boundaries between the machine, in which everything is running backwards, and the normal world. I think I can tell him. If we suppose that our time-travellers shoot a deer with a bow and arrow somewhere in the vicinity of Sherwood Forest, a diagram of the course of their arrow would look like this:



Celine Denruyter wrote a delightful story, which, together with Mr. Godfrey-Smith's essay, I recommend be published. These two are my runners up. In her feminist version of my story it is Jocasta who travels, and Dum-Dee who stays behind. Miss Denruyter's story goes one better than mine in that, where I have only the Dum-Dee person, she has the Jocasta (mother)—Jocasta (daughter)—Kalbu—Kalba person (or, more likely, since women are animals, animal). And if the life-line of Dum-Dee is (as it is) like a paper clip, the more complicated life-line of her (predominantly female) heroine is like the hair spring of a watch. (She could not have known that in an earlier version of my story, the instructions for building the time machine were written on an egg, which was laid by a hen which hatched from the egg.) I could fault her story only in that she does not explain how the dog who leads Jocasta (mother) to the deep freeze is called Kalba, while the dog who is taken on the journey and when mother and daughter eat, is called Kalbu. If Kalba has a mother, I suppose it must—on the principle that, if you change into a dog your mother becomes the dog's mother—be Jocasta herself. The story is well worth carefully working out.

I have awarded the prize to Margarita Levin. She does not admit to seeing many of the difficulties with the story, and she is excessively phlegmatic about the arrival *ex nihilo* of a machine containing one eccentric human being and the remains of another. She also swallows rather uncritically the distinction between personal and external time. I think it is better to regard these alleged time-travellings as cases when whole cross-sections in the life of a human being, mental and physical, have got displaced (though this does not explain the travelling part, and David Lewis dispenses with it). But Mrs. Levin's views are the most level headed and carefully logical and, though her diagram is a bit inelegant, she is the only competitor except Miss Denruyter to have worked everything out correctly—better than I had myself in fact, for I had not realised that Jocasta, poor woman, could commit incest by only one act of sexual intercourse.[1] Since two competitors have accused me of sexism, one overtly and one by implication, may I point out that both the winner and one of the runners up are women?

---

[1] If you were to define 'incest' as having sexual intercourse with a person who was the product of your own previous act of sexual intercourse, then Jocasta *would*, so to speak, have committed a logically impossible crime.

University of Nottingham

SWORDS' POINTS

By MARGARITA R. LEVIN

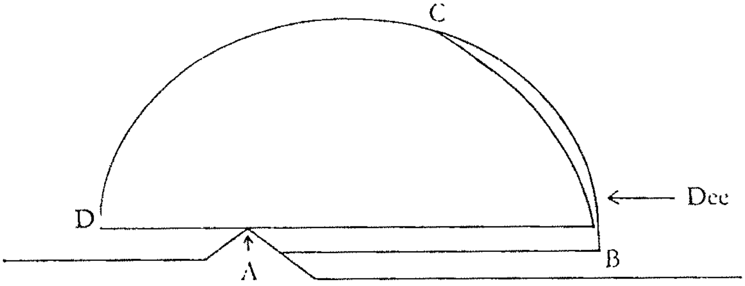
YES, poor Jocasta is unwittingly guilty of incest—she had sexual relations with her own child. Jocasta's crime is remarkable, however, because it demonstrates that only one act of intercourse is necessary for a woman to be guilty of incest. Ordinarily, two are performed: the first being the act that results in the child's conception, the second being union with the child himself. What happens in Jocasta's case is that the father of her son is a later stage (from his point of view) of that very son. So only the single act of intercourse during which Dee was conceived sufficed for Jocasta's transgression.

Now it is customary for the stages of a man's life to succeed one another in external time as they do in personal time (the phrases are David Lewis', 'The Paradoxes of Time Travel' *American Philosophical Quarterly* Vol. 13, No. 2, (April 1976), p. 146). The date on which he is born always precedes the date on which he marries and the date on which he dies. However, it is not logically necessary that one's life be linear in external time. There is no literal contradiction in the story of

Dum, who fathers himself and is devoured by himself. After all, at no point in the story is the same stage of his life obliged to be in two places at once, or both to have and not have any other property.

To be sure, we are puzzled by the loop in time that Dum's life describes: how did it get started and who wrote the book about building a time machine? However, these questions are no different from questions about where *anything* originally came from. We can ask about the origin of the atoms that make up Jocasta; their time line is not neatly presented to us. The atoms either go back endlessly, or if the universe is finite, they just start. In either case the question of ultimate origin is as unanswerable as the question of the book's origin. What makes us think that when such questions are asked about the loop they are different and *ought* to be answerable is that the entire loop is open to inspection. *Sub specie aeternitatis* this difference disappears.

The following diagram shows the time lines of Dee, Dum and Jocasta.



A marks Dee's conception, the point at which he separates from his later stage, which is called 'Dum'. B is the point at which Dee and his later stage Dum set off in their time machine. C is the point at which Dee eats Dum; the last stage of Dum merges with that stage of Dee. D is the point at which our hero resumes life in accordance with external time. For convenience, let us say that this is also the point at which he changes his name to Dum. The lowest line which never loops is Jocasta's time line.

A then is the point at which Jocasta committed her heinous crime—mating with her son. As mentioned earlier, it might have been thought to be analytic that incest requires two acts of intercourse. But, as Quine notes in another context, 'science fiction and analyticity are at swords points' (W. V. O. Quine, review of *Identity and Individuation*, Milton K. Munitz, ed. in *Journal of Philosophy*, Vol. LXIX No. 16, September 7, 1972, p. 496).

University of Minnesota

© Margarita R. Levin 1980

JOCASTA'S CRIME: A SCIENCE-FICTION REPLY

By CELINE DENRUYTER

'THOUGH reading science fiction is one excellent way of acquiring the necessary imaginitive insight, writing it is, I dare say, even better' (J. Harrison, 'Dr. Who and the philosophers', p. 24). At any rate, I have made an attempt.

'If you make any counterfactual supposition and hold all else fixed you get a contradiction' (D. Lewis, 'The paradoxes of time travel, p. 152). So I had to change everything. . . .

In the late (latest) Dum Jones's files the following letter was discovered:

I was sitting solidly frozen in my deep freeze when suddenly it was opened by a woman who called herself Miss Jones and who had been led to me by her dog. I fell in love with her and married her. After a decent interval we produced a baby whom, because she was the spitting image of her mother, we forenamed Jocasta. When daughter Jocasta reached oedipal maturity, she produced a wild attack of paranoia. With the help of the learned dog Kalbu—who possessed the knowledge of building time machines— she built a time machine and got inside, taking her mother and Kalbu; and they started travelling backwards. Because of food shortage they were compelled to kill and eat Kalbu, but mother Jocasta could not digest the canine food and died after a few days. Subsequently she underwent the not uncommon (Plato, *Phaedrus* 249 a–b) transformation into a learned dog, which daughter Jocasta called Kalba, because she was the spitting image of the deeply regretted Kalbu. Daughter Jocasta finally got out of the machine together with Kalba and she blew up the time machine. She was prudent enough to keep Kalba with her, in case she needed to build another time machine. Jocasta, however, was not happy in her old environment. She wandered about and tried to remember the future things that had happened during the time travel. The next thing she remembered was being led by Kalba to a rusty deep freeze . . .

Did I (who am I?) commit a logically possible crime by sitting in that deep freeze?

© CELINE DENRUYTER 1980

TRAVELLING IN TIME

By William Godfrey-Smith

MISS JOCASTA'S actions [1] not only fall below the standards of accepted social conventions; they are also biologically, physically and metaphysically questionable.

'Travel' in time, *prima facie*, must be either into the future or the past. The former (in one sense) is trivial and the latter (in any sense) is absurd. A journey into the past presupposes that past 'times' are still 'there', like hidden places; elusive, perhaps empirically inaccessible, but *conceivable* and so—Professor Harrison ([2], p. 24) would have us believe —*logically possible* destinations to which one might *travel*. The blunt objection to time travel is that there is just nowhere to *go*. Time travel talk is the result of literally interpreting the spatial metaphors which we employ ubiquitously in describing our temporal experience.

The metaphysical picture which underlies time travel talk is that of the block universe, in which the world is conceived as extended in time as it is in space. This is a preposterous (if perennially seductive) metaphysical conception which is not easily demolished; perhaps Miss Jocasta's predicament may help to knock a few more chips from its foundations.

Forward travel may be dismissed as boring. Backwards travel must be dismissed as contradictory; it involves the absurdity of supposing that some event is both earlier and later than another. To escape this contradiction the friends of time travel find it necessary to introduce additional time series, distinguishing 'objective' and 'subjective' time ([2], p. 10) or 'external' and 'personal' time ([3], p. 146). This intolerably mystifying temporal dualism—quite devoid of any empirical foundation —is never satisfactorily explained. (The plurality of 'local proper times' in physical theory in no way sanctions these bizarre backwards chronologies. Quite the contrary!)

The biological problem is the following. Dee is the son of Dum and Jocasta. So Dee obtained half his genes from Dum and half from Jocasta. But Dum is diachronically identical with Dee, and is therefore genotypically identical with him (*i.e.* himself). That is, Dee is both genotypically identical with and distinct from Dum, which is absurd. (And what if, instead of Dee, Dum and Jocasta had had the girl which they were *really* hoping for?)

The physical problems are no less pressing. Let us ponder what happens during the journey. For the backwards traveller no sun would be visible rising in the West (*cf.* [2], pp. 11–13); radiant energy would diffuse upward from the earth to the sun. The surroundings of the machine would be invisible. Moreover the causal interactions at the boundary of the backwards travelling machine are not easy to explain. It is of course easy to imagine physical processes being reversed—clocks running backwards, eggs unscrambling, and so forth. What is hard to imagine, indeed impossible to make sense of, is backwards *continuing*—later states being earlier than the states they are later than—which is what a coherent description of the journey requires. Furthermore, if Feynman is right, and particles travelling backwards in time are antiparticles, any interactions at the boundary of the travelling machine will be of the most explosive kind!

Metaphysically, Dum's and Dee's journey presupposes the reality of the future in a very strong sense. It is a necessary condition for Dum's existence in AD 1980 (say) that he has *already existed* in AD 2000 (say). Backwards travelling assumes that there is a real and determinate future which is Dee's (Dum's) starting point—the point of embarkation of his journey in time. If this is true then the future is fixed and determinate (even if *undetermined physically*), and a *fatalistic* (though not necessarily *deterministic*) metaphysic will inevitably ensue.

But if Jocasta's predicament presupposes the truth of fatalism then she may reflect—indeed *must* reflect given that she reflects at all—that her 'crime', if logically *possible* is logically *necessary*. (Miss Jocasta, we may note, was indeed much disposed to philosophical reflection: contrary to the suppositions of male chauvinist professors ([1], p. 65) her lack of interest in engineering did not bespeak an exclusive preoccupation with men.) So her 'crime', if crime it be, is one for which she was not after all morally culpable. We are not after all to blame for the necessary truths of the world.

Australian National University © WILLIAM GODFREY-SMITH 1980

REFERENCES

[1] Jonathan Harrison, 'ANALYSIS "Problem" No. 18', ANALYSIS, March 1979, pp. 65–6.
[2] Jonathan Harrison, 'Dr Who and the Philosophers', *Proceedings of the Aristotelian Society, Supplementary Volume* 45, 1971, pp. 1–24.
[3] David Lewis, 'The Paradoxes of Time Travel', *American Philosophical Quarterly* 13, 1976, pp. 145–152.

SYNTACTICAL TREATMENT OF NEGATION

By ROBERT K. MEYER

IN what has sometimes been taken as a body blow to syntactical interpretation of modal logic, Montague showed in [1] that, under certain apparently plausible conditions, inconsistency is bound to result when necessity is taken as definable within any classical theory that contains a modicum of formal arithmetic. From this [1] draws such conclusions as 'Thus if necessity is to be treated syntactically, . . . then virtually all of modal logic, even the weak system S1, must be sacrificed.' (p. 294) And also, 'Thus it seems at present doubtful that any philosophical interest can be attached to S1–S4.' (*Ibid.*)

These are strong conclusions. It must, accordingly, be seen as a pity that the premisses on which they are drawn are totally implausible. (How in fact modal logics should be syntactically interpreted will be made plain in [3], along lines already presaged in [2].) For what Montague assumes is that to interpret necessity syntactically is to interpret it as a (proof) predicate of Gödel numbers of expressions. Let Nx be such a one-place predicate, and suppose that we have enumerated the formulas of a system S effectively, i.e. assigned them Gödel numbers, using A^I for the Ith formula in the enumeration and letting I be its Gödel number. Then, for familiar theories containing enough arithmetic, it is well known that we can find a formula NX, with sole free variable X, such that, for each number I, and corresponding numeral ι in the system S, we have

$$(1) \quad \vdash_s N\iota \text{ iff } \vdash_s A^I$$

So the proposal which Montague considers is that we should pick some NX which satisfies (1) for all A^I, and take it as a syntactical counterpart of necessity for the system S. So parsed, necessity is of course a *predicate*, not a *connective*, a fact which [1] alleges to have certain philosophical advantages pointed out by Quine and Carnap. Even as a predicate however, this defined necessity may be expected to satisfy some of the laws of Lewis-style modal logics. Alas, as [1] shows, it can't satisfy together a minimal plausible collection of these laws, whence the disturbing conclusions cited above are drawn.

However, the option that [1] does not consider is that the fly in the ointment lies *not* in identifying necessity with provability (whether *predicatively* or *connectively*, a question which need concern us here no longer), but with the *ins* and *outs* of arithmetizing syntax itself. The problem in doing so, as we all know, is that our language gets all tangled up in itself, whence we can turn it on itself to produce Gödelian paradoxes and near paradoxes. To clinch the point, I shall argue, along lines reminiscent of [1], that a syntactical treatment of negation is impossible, too. Specifically, there is no way to formalize, within the sorts of systems considered in [1], the *predicate* of sentences 'is not the case' so that, arithmetized, every statement $\sim A$ turns out to be provably materially equivalent to a statement that 'A' is not the case.

This, it might be thought, is merely an instance of the Tarskian point, from [4], that arithmetic truth is not arithmetically definable. Technically, in some sense, it is. But philosophically it is not; for we are not imposing on our negation-predicate any substantive conditions that require it to say that 'A' is false, for arbitrary A. We are merely requiring that, in our system S, we should have a theorem, for each given sentence A,

(2) $\sim A \longleftrightarrow$ 'A' is not the case

Since, normally, we take the two sides of (2) not merely as materially equivalent but as actually synonymous, there are only two conclusions that we can possibly draw: either (a) there is something wrong with '$\sim$' as a formal counterpart of the intuitive connective 'not', or (b) there is something wrong with looking to arithmetization as a way of providing formal counterparts to clear intuitive connections between language and metalanguage.

In order to make this point as plain as possible, let us begin not with an effort to *define* 'is not the case' within some antecedently given system; rather, following our intuitions, let us simply take '$-$' as a 1-place primitive *predicate*, to be added to the *language* of our antecedently given system S. If the old language was L, let the new language be L$-$. And now let us get a system S$-$ by assigning Gödel numbers to the formulas of L$-$, and taking as new *axioms* each *sentence* (i.e., closed formula) of L$-$ of the form

(3) $\sim A^I \longleftrightarrow -\iota$

where ι is again a numeral corresponding to the number I. (3), evidently, is a formal counterpart of (2), which it is intended to express. And now let S$-$ be the classical first-order theory got by taking as axioms all theorems of S and all instances of (3).

We have not yet said much about S, though we have made some presuppositions about it (e.g., its language L contains some terms that will do as *numerals*, which the deductive structure of S allows us to take as *names of numbers*). In fact, all that we require of S, in addition, is that every recursive function be strongly definable therein. E.g., S could be first-order Peano arithmetic P, or the system N of all first-order arithmetical truths in one of the usual arithmetical languages, or Robinson's finitely axiomatized subsystem Q of P. It could also, of course, be some more comprehensive system in which arithmetical entities are definable, like the set theory ZF.

Since S— is an extension of S, all recursive functions are strongly definable there also. This permits proof of the following *lemma of reflexive reference* (so-called by Dwyer, though an old story that may be gleaned, e.g., from [1]). Let BX be a formula of L—, in which X may occur free. Then for some formula A^I, we have

(4) $\vdash_{S-} A^I \longleftrightarrow B\iota$

where $B\iota$ is the result of proper substitution of the numeral ι corresponding to the number I for free X in BX. This result holds for all formulas BX of L—, and A^I may be taken to be a sentence if no variables other than X are free in BX. In particular, $-X$ is such a formula, and so we have, for some sentence A^I with Gödel number I and corresponding numeral ι,

(5) $\vdash_{S-} A^I \longleftrightarrow -\iota$

But, we also have, as an instance of our axiom (3) of S—,

(6) $\vdash_{S-} \sim A^I \longleftrightarrow -\iota$

for this particular I, whence moreover, truth-functionally,

(7) $\vdash_{S-} A^I \longleftrightarrow \sim A^I$

I.e., S— is inconsistent.

The key step in this argument is of course (4), which corresponds to the key step in [1]. And, while we have carried out this argument for an extension of S, we could of course have carried it out for S itself in case there is any formula $-X$ already in L for which all instances of (3) are already provable in S.

So we are stuck with our dilemma between (a) and (b) above. We should, of course, dislike the conclusion that there is something wrong with our connective '$\sim$'. But this conclusion should not be dismissed too hastily. For many of the anomalies which afflict familiar formal systems—not excluding Gödelian anomalies—seem to cluster around this connective. (It may be recalled that Anderson, in [5], presented a tongue-in-cheek attack on '$\sim$', simply by transposing familiar attacks on modal connectives (e.g., they confuse use and mention, as is indicated by circumlocutions like 'it is necessary that'; Anderson points out that 'it is not the case that' is a similar circumlocution, not easily avoided when we are trying to speak grammatical English.) Perhaps, however, the attack should not have been so tongue-in-cheek. For, in many respects, 'not' makes a better predicate than it does a connective—maybe grammatical English is trying to tell us something.)

However, conclusion (b) is the one that is more likely to be drawn:

namely, that the method of arithmetization is *not* the way to pursue a negation-predicate, for that pursuit will turn reflexively upon itself to our chagrin. Again, note that this pursuit is ruled out in a very strong sense; not only can't we *define* a negation predicate, but we can't even consistently add such a predicate while obeying the fair-and-square rules of arithmetization. To remind the reader: it is not a falsity predicate for which we are looking, which would perforce divide the sentences between truths and falsehoods. It is only a negation predicate, which can be as incomplete as we please—indeed, as incomplete as our original negation connective is incomplete, since it is exactly that connective which our negation predicate is supposed to reflect.

Now let us note one more thing. There would be no problem about *adding* a negation predicate if we restricted our axioms (3) to formulas $\sim A^I$ that *already* belong to the *original* language L; or, equivalently, if we restricted our enumeration of formulas to formulas of the original language L. For as inspection of any standard proof will reveal, it is essential in proving the lemma of reflexive reference that A^I be permitted to contain the predicate B. Equivalently, there would be no problem if we restricted our enumeration of formulas to formulas of the original language L. (We might then, for example, simply interpret $-I$ to mean that the formula A^I is false in the standard model, whence it is clear that, for each choice of S contained in the system N, all axioms of *restricted* systems S— presently being considered will be true in the standard model, whence such S— will be consistent.) This brings out even more clearly that the *inconsistency* of our *unrestricted* systems S— reveals no difficulty in the *notion* of a negation-predicate; again, the difficulty lies in trying to arithmetize too many things at once.

We may now return to our original concern: namely, an assessment of the import of [1] for syntactical treatments of modality. (This is, of course, our real concern, since otherwise the main results of this paper must be taken as known, even if not sufficiently appreciated or understood in the proper context). The whole burden of the philosophically interesting part of the argument (as it applies to modal logic, since [1] contains other technically interesting results not marred by philosophical misapplication) lies in the demand that we shall take, as our formal counterpart of 'it is necessary that', some predicate NX that is *already definable* within the system S, and which, without restriction, satisfies the additional plausibility criteria set out there.

Suppose that S were already a modal theory, equipped with a primitive connective $\Box$ as we have assumed S equipped with a primitive $\sim$. What then would we make of the results of [1]? As much, presumably, and no more, as we have made of the similar situation with $\sim$; namely, that we cannot define within S a necessity-predicate that plausibly corresponds to our necessity-connective. But there is no bar against *adding* such a predicate, satisfying all the plausibility conditions laid out in [1], so long as we do not also insist that it feed cannibalistically upon itself by supposing that our necessity-predicate will mean provability-in-the-extended-system. It suffices, after all, for our syntactical interpretation of modality that it should mean provability-in-the-original-system. And here, *nihil obstat*.

A final word is here in order. This paper, to say the least, objects to [1]. But the objection, in the end, is not so much an objection to technique as to tone. For it is the tone which produces the errors, and which has discouraged the pursuit of fruitful lines of inquiry. It has been abroad for years—due, no doubt, above all to the pernicious influence of Quine in this subject—that some syntactical anomaly, some confusion, has guided the development of modal logic from the start, from which the subject might be rescued only by a satisfying extensional interpretation. [1] sharpens this sense of anomaly. But it is a dubious sharpening, since the same arguments apply, *mutatis mutandis*, to pose difficulties for extensional interpretation of presumptively extensional notions. In interpreting $\Box A$, in many contexts, 'it is provable that A' looks like a favoured choice; but, certainly, in interpreting $\sim A$, 'it is not the case that A' is the favoured choice. And, evidently, any argument that rules out the former on grounds that may be equally well applied to rule out the latter must be rejected as philosophically spurious. He who finds modal motes in his brother's eye had better first remove the log of Boolean negation from his own.

Australian National University © ROBERT K. MEYER 1980

REFERENCES

[1] R. Montague, 'Syntactical treatments of modality, with corollaries on reflexion principles and finite axiomatizability', *Acta Philosophica Fennica*, 16 (1963), pp. 153–67 (reprinted in R. Montague, *Formal Philosophy*, ed. R. H. Thomason, Yale, 1974, pp. 286–302; page numbers here come from the latter).

[2] R. K. Meyer, 'Coherence in modal logics', *Logique et Analyse*, 14 (1971), 658–668.

[3] R. K. Meyer, *Coherence Revisited*, in preparation.

[4] A. Tarski, *Logic, Semantics, Metamathematics*, ed. J. Woodger, Oxford, 1956.

[5] A. R. Anderson, 'Does negation rest on a mistake?', *Philosophical Studies*, 16 (1965), 4–6.

WHAT DO THE CONSISTENCY PROOFS FOR NON-EUCLIDEAN GEOMETRIES PROVE?

By GEOFFREY HUNTER

DESPITE the good work of Raymond Bradley (1964) and Stephen Barker (1964 ch. 3 and 1967) friends still regularly draw the old familiar invalid inferences from the existence of non-Euclidean geometries; so it may be worth trying to show yet again that the first two of the propositions cited below follow from known mathematical results (about the interpretability of non-Euclidean geometries in terms of Euclidean) only when taken in a strictly qualified sense, and that the third one does not follow at all:

(1) 'If Euclidean geometry is consistent so is Lobachevsky's/Riemann's.'
(2) 'The negation of Euclid's Parallel Postulate is not contradictory.'
(3) 'Euclid's Parallel Postulate is not a necessary truth.'

It will make the essentials of my argument clearer if I simplify a situation that is in reality a good deal more complicated and take Euclidean geometry to be given by the five postulates:—

1. To draw a straight line from any point to any point.
2. To produce a finite straight line continuously in a straight line.
3. To describe a circle with any centre and distance.
4. That all right angles are equal to one another.
5. Through a point not on a given straight line there is one and only one straight line parallel to the given line.

(The first four are from Sir Thomas Heath's translation of Euclid. The fifth is a substitute for Euclid's own fifth postulate and is commonly called "Playfair's Axiom", although it was stated long before Playfair, by Proclus in the fifth century A.D.)

Again simplifying a more complicated reality, we may take Lobachevskian (hyperbolic) geometry to be given by Postulates 1–4 and the replacement of 5 by

> 5L: Through a point not on a given straight line there is *more than one* straight line parallel to the given line

and Riemannian (elliptic) geometry to be given by Postulates 1–4 and the replacement of 5 by

> 5R: Through a point not on a given straight line there is *no* straight line parallel to the given line.

Then the standard relative consistency proofs for the non-Euclidean geometries proceed by showing that if key terms (e.g. 'point', 'straight line', 'circle', 'distance', 'angle') of the non-Euclidean postulates are interpreted in suitable ways the postulates so interpreted will all come out as postulates or theorems of Euclidean geometry. Hence if a pair of contradictory theorems were provable in a non-Euclidean geometry then a pair of contradictory theorems would be provable in Euclidean; and that gets summed up as 'If Euclidean geometry is consistent so are the non-Euclidean geometries'.

So, for example, if in the Lobachevskian postulates (1–4 and 5L) we interpret 'point' as 'point inside a fixed (Euclidean) circle C', 'straight line' as 'open chord [i.e. omitting endpoints] of the circle C', and give a suitably complicated definition of 'distance' (e.g. distance $PQ = \log ((QO . PS) \div (PO . QS))$ where OPQS is a chord of the circle C with O and S on the circumference), the postulates, so re-interpreted, all come out as postulates or theorems of Euclidean geometry. Now suppose Lobachevsky's system is inconsistent in the sense that, for some sentence A, both A and $\sim$A are theorems of Lobachevsky. Interpret A according to our dictionary: let the result be A'. Then A' and $\sim(A')$ will both be theorems of Euclidean geometry. So if Lobachevsky's system is formally inconsistent then so is Euclid's. So if Euclid's system is formally consistent so is Lobachevsky's. Similarly, if in the Riemannian postulates (1–4 and 5R) we interpret 'point' as 'point on the surface of a fixed (Euclidean) sphere S' and 'straight line' as 'great circle of S' (allowing retracement), and take the distance between two points as the distance along a great circle containing them (allowing retracement), the postulates come out as theorems of the Euclidean geometry of a sphere. And so again if Euclid's system is formally consistent so is Riemann's.

At this point (1), (2) and (3) are often drawn as conclusions. To see whether these conclusions are rightly drawn let us consider the following comparable case:

Postulates for a fragment (N) of Natural Number Theory

1. Every natural number greater than 1 has a successor.
2. Every NN greater than 1 is greater than at least two NNs.
3. Every NN greater than 1 has at least one immediate predecessor.
4. Every NN greater than 1 is either odd or even.
5. Every NN greater than 1 has exactly one immediate predecessor.

Postulates for Unnatural Number Theory (U)

1–4: as for N above.

5′. Every NN greater than 1 has at least two immediate predecessors.

Is U consistent?

Standard argument:—

'Interpret "has an immediate predecessor" as "is greater than some natural number".
Then Postulates 1–4 and 5′, *so interpreted*, are theorems of N.
Conclusions:

(1) So if N is consistent so is U.
(2) So the negation of Postulate 5 of N is not contradictory.
(3) So Postulate 5 of N is not a necessary truth.'

What this example shows is that, if conclusions (1) and (2) can be validly drawn, then the expressions 'consistent' and 'not contradictory' that occur in them do *not* imply 'free from conceptual absurdity', since 5′ is not free from conceptual absurdity. Further, conclusion (3) cannot be drawn at all, since 5 *is* a necessary truth. Applying this result to the case of Euclidean and non-Euclidean geometries we have, similarly:

(1) The relative consistency proofs for non-Euclidean geometries do *not* prove consistency *where that implies freedom from conceptual absurdity.*
(2) They do *not* prove that the negation of Euclid's Parallel Postulate is not contradictory *where that implies 'free from conceptual absurdity'.*
(3) They do *not* prove that Euclid's Parallel Postulate is not a necessary truth[1] (and so don't show that genuine contraries of it are not necessarily false.)

What do they prove? The answer is:

> They prove that if Euclidean geometry is consistent, then the *formal* systems L′ and R′ are *formally* consistent, where L′ and R′ are the results of replacing the non-logical expressions in Lobachevsky's and Riemann's geometries by schematic letters. That is all.

This leaves it open that Euclidean geometry, understood as Euclid intended, is indeed a system of necessary truths. The problem is: what did Euclid intend? It is perfectly clear that he did not intend his *Elements* to be simply an uninterpreted system. In the first place, the very idea of a formal system, i.e. something definable in terms of strings of marks without any reference to meaning, is not to be found until thousands of years later. Secondly, Euclid prefaces his system with 23 definitions, many of which, including his definitions of 'point', 'line' and 'straight line', are elucidations of terms that occur in his system by means of terms that do not, and (unlike the definitions in modern formal systems) are never made use of in the ensuing system. So why did he put them in? If we look at his definition of 'straight line' I think we can see why:

---

[1] For a similar reason, proofs of the independence of, say, the Axiom of Choice from the other axioms of set theory do not show that the Axiom of Choice is not a necessary truth.

> A straight line is a line which lies evenly with the points on itself.

Here he is very carefully *not* defining 'straight' in terms of paths of light rays or the paths taken by taut plumb lines or anything like that, but instead deliberately trying to avoid any such interpretations. In Euclid's *Optics*, as Heath points out (vol. I p. 166), it is a *hypothesis*, and *not* a matter of definition, that opseis (lines of sight) are straight. Heath cites pre-Euclidean definitions of 'straight' that appeal implicitly to the notion of *line of sight*, and concludes: 'We may safely infer that Euclid's definition was a new departure of his own'. I conclude that Euclid is quite deliberately trying to define 'straight' without reference to light rays, and that interpretations of his axioms in terms of the paths of light rays (or plumb lines) not only are not Euclidean but are deliberately rejected by Euclid. Euclid's Parallel Postulate is *not* an empirical hypothesis about, e.g., laser beams, and nothing a laser beam does can refute it. I take it that what Euclid is claiming is that there is a notion of 'straight line' which is not to be identified with the notion of 'path of a light ray in a homogeneous medium' or 'path of a taut plumb line', and which is such that it is a contingent matter whether or not the path of a light ray in a homogeneous medium *is* a straight line; that whatever empirical aids we use for testing for straightness we employ because we have an intellectual idea of what it is for something to be a straight line (viz. a line that 'lies evenly with the points on itself', Euclid's attempt to catch the notion in words, which is something I cannot improve on).

Einstein's (1921) division of respectable geometries into just two sorts, (1) 'purely axiomatic' (meaning 'wholly uninterpreted') and (2) 'practical' (i.e. containing 'affirmations as to the behaviour of practically-rigid bodies', affirmations that 'rest essentially on induction from experience'), does not exhaust the possibilities, and in particular does not cover the present interpretation of Euclid (which I think he would have characterized as mere 'mystic obscurity'). When physicists debate whether space is Euclidean or Riemannian, they are talking about reinterpretations of Euclid, not Euclid. Of course they may say, 'If that is what Euclid is talking about, then we are not interested in it: it is of no importance to physicists'. That is up to them. But saying 'I am not interested in numbers but in the conditions in which water-drops coalesce' does not refute number theory: and a physicist who finds, e.g., that a laser beam projected into space returns after many years to hit him in the back of the neck does not have there material with which to refute *what Euclid meant.*

University College of North Wales, Bangor © GEOFFREY HUNTER, 1980

REFERENCES

Stephen F. Barker, *Philosophy of Mathematics,* Prentice-Hall, Englewood Cliffs, 1964.
Stephen F. Barker, 'Geometry', in Paul Edwards, ed., *The Encyclopedia of Philosophy,* Macmillan, N.Y., 1967, vol. 3, pp. 285–290.
Raymond D. Bradley, 'Geometry and Necessary Truth', *Philosophical Review,* 73 (1964), pp. 59–75.
Albert Einstein, 'Geometry and Experience' (1921), in his *Sidelights on Relativity,* Methuen, London, 1922.
Sir Thomas L. Heath, *The Thirteen Books of Euclid's Elements,* 2nd edition, Dover, N.Y., 1956.

FREGE ON 'POSSIBLE CONTENT OF JUDGMENT'

By RONALD E. NUSENOFF

IN both 'Über Begriff und Gegenstand' (1892) and the introduction to Volume I of *Grundgesetze der Arithmetik* (1893), Frege notes a change in his views from *Begriffsschrift* (1879) and *Die Grundlagen der Arithmetik* (1884).

> Formerly I distinguished two components in that whose external form is a declarative sentence: (1) the acknowledgment of truth, (2) the content that is acknowledged to be true. The content I called a 'possible content of judgment'. This last has now split for me into what I call 'thought' and 'truth-value', as a consequence of distinguishing between sense and denotation of a sign.[1]

I find this explanation of 'possible content of judgment' (beurtheilbarer Inhalt) puzzling. In the *Begriffsschrift,* a possible content of judgment seems to be what a sentence is used to assert, i.e., a thought (Gedanke) rather than some combination of thought plus truth value. Frege's 1910 letter to Jourdain supports the identification of possible contents of judgment with thoughts as the senses of sentences: 'Instead of "beurtheilbarer Inhalt" we can also say "Gedanke"'.[2] Further, Frege did not reify truth values as objects until after the *Grundlagen* when, upon making the distinction between sense and reference, and taking sentences to be a type of name, he needed some objects to serve as the referents of sentences.

---

[1] *Grundgesetze der Arithmetik,* trans. by M. Furth as *The Basic Laws of Arithmetic* (Berkeley and Los Angeles, 1964), pp. 6–7; cf. p. 38 fn., and 'Über Begriff und Gegenstand', trans. as 'On Concept and Object', in Geach and Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege* (Oxford, 1952), 42–55, p. 47.

[2] Cited in J. van Heijenoort, *Frege and Godel: Two Fundamental Texts in Mathematical Logic* (Cambridge, 1970), p. 11 fn. This line was *not* cited in Jourdain's 'The Development of the Theories of Mathematical Logic and the Principles of Mathematics', *The Quarterly Journal of Pure and Applied Mathematics* 43 (1912), 219–314.

The content of a predicate also seems roughly equivalent to what Frege later took to be its sense, this being distinct from its reference, which he took to be a concept. However, the content of a name is clearly identifiable with what Frege later took to be its reference.[3] In the *Begriffsschrift*, Frege took identity to be a relation, between names, of having-the-same-content (Geach and Black, pp. 10–12). In 'Über Sinn und Bedeutung' (1892) Frege accounts for the cognitive value of informative identity sentences in terms of 'a difference in the mode of presentation of that which is designated', with the sense of a sign being 'wherein the mode of presentation is contained' (Geach and Black, p. 57). In the *Begriffsschrift*, the names are said to 'correspond to', or, 'go along with', 'ways of determining content' (Geach and Black, pp. 11–12). Hence in the case of (nonsentential) names, content is no more than reference, and sense is combined with sign (if with anything at all).

Why did Frege claim that both sense and reference had been combined within the notion of possible content of judgment? Perhaps he was reading his later characterization of sentences as names, plus the principle that the reference of a sentence is a function of the referents of its constituents, back into his earlier works. Then the contents of the constituents of a sentence would determine a truth value as part of the content of that sentence.

This explanation does not seem compatible with the *Grundlagen*. After claiming fidelity to the principles,

> (1) always to separate sharply the psychological from the logical, the subjective from the objective;
> (2) never to ask for the meaning of a word in isolation, but only in the context of a proposition;

Frege goes on to explain,

> If the second principle is not observed, one is almost forced to take as the meanings of words mental pictures or acts of the individual mind, and so to offend against the first principle as well.[4]

It was only *after* Frege formulated the notion of sense, as an objective notion of meaning, that he could see how to ask for the meaning of a

---

[3] Thus I agree with Michael Dummett's explanation of 'content': '(the word "content" being taken in a vague sense as roughly covering both what Frege later called "sense" and what he later called "reference")', *Frege, Philosophy of Language* (London, 1973), p. 315. Dummett, however, does not notice the problem with Frege's claim that 'possible content of judgment' covered truth values as referents.

[4] *Die Grundlagen der Arithmetik*, trans. by J. L. Austin as *The Foundations of Arithmetic* (Oxford, 1950), p. x; cf. Section 60.

name is isolation without offending against the first principle. Prior to replacing the notion of content with the notions of sense and reference, Frege had been unable to provide a way of breaking down the content of a sentence which succceeded in relating the content of a sentence to the contents of every one of its constituents. After having made this replacement, Frege no longer needed to emphasize the second principle.

The distinction between sense and reference may thus have been not merely a useful alteration to the theory of the *Grundlagen*, but in fact a necessary one. Frege's realization of this necessity, along with his characterization of sentences as names, would *explain* his claim that possible content of judgment had been a combination of both thought and truth value. Given that neither the realization nor the characterization is evident in the *Grundlagen*,[5] Frege's claim looks to be an explanation of what he later thought he should have meant, rather than of what he had at that time meant.

California State University, Northridge

© RONALD E. NUSENOFF 1980

---

[5] Austin notes at the bottom of p. 1 of the *Grundlagen*: '(I have tried throughout to translate *Bedeutung* and its cognates by "meaning" and *Sinn* and its cognates by "sense", in view of the importance Frege later attached to the distinction. But it is quite evident that he attached no special significance to the words at this period.)'

AN EXTENSION OF A PROOF OF PRIOR'S

OR

WHEN THINKING MAKES IT SO

By J. J. MacIntosh

PRIOR has a proof that, where D is any operator such that if α is a wff, so is $D\alpha$, then:

$$D\exists p(Dp \,\&\, {\sim}p) \rightarrow \exists p(Dp \,\&\, {\sim}p \,\&\, p \neq \exists p(Dp \,\&\, {\sim}p))$$

(A. N. Prior, *Objects of Thought*, Oxford 1971, pp. 86–7; there are minor misprints in the proof—lines A, C and G—in the 1971 edition).

This result can be generalized. For convenience I introduce the following notation:

M: $M^np =^{df} p = \exists pF^np$

F: $F^1p =^{df} Dp \,\&\, {\sim}p$

and, for $n \geqslant 2$, $F^np =^{df} F^{n-1}p \,\&\, {\sim}M^{n-1}p$

With this notation, Prior has proved

$$D\exists pF^1p \rightarrow \exists p(F^1p \,\&\, {\sim}M^1p)$$

i.e., $D\exists pF^1p \rightarrow \exists pF^2p$

To generalize this result we first prove the lemma:

$$\exists pF^np \longleftrightarrow \exists pF^{n+1}p,\ n \geqslant 1$$

Proof:

1	(1)	$\exists pF^np$	A
1	(2)	$\exists p(F^np \,\&\, M^np) \vee \exists p(F^np \,\&\, {\sim}M^np)$	1, taut
3	(3)	$\exists p(F^np \,\&\, M^np)$	A
3	(4)	$\exists p(F^np \,\&\, p = \exists pF^np)$	3, M
3	(5)	$F^n\exists pF^np$	4, E.I.
3	(6)	$D\exists pF^np \,\&\, {\sim}\exists pF^np \,\&\, {\sim}M^1\exists pF^np \,\&\, \ldots \,\&\, {\sim}M^{n-1}\exists pF^np$	5, F
3	(7)	${\sim}\exists pF^np$	6, &—
3	(8)	$\forall p{\sim}F^np$	7, Q E.
3	(9)	${\sim}F^n\exists pF^np$	8, U.I.
	(10)	${\sim}\exists p(F^np \,\&\, M^np)$	3, 5, 9, RAA

1 (11) $\exists p(F^n p \;\&\; {\sim}M^n p)$ 2, 10, taut
1 (12) $\exists pF^{n+1}p$ 11, F
13 (13) $\exists pF^{n+1}p$ A
13 (14) $\exists p(F^n p \;\&\; {\sim}M^n p)$ 13, F
13 (15) $\exists pF^n p$ 14, taut

We are now in a position to derive the following generalization of Prior's result:

$D\,\exists pF^n p \rightarrow \exists pF^m p$ for any $n, m \geqslant 1$

Proof:

1 (1) $D\,\exists pF^n p$ A
2 (2) ${\sim}\,\exists pF^n p$ A
1,2 (3) $D\,\exists pF^n p \;\&\; {\sim}\,\exists pF^n p$ 1, 2, &+
1,2 (4) $\exists p(Dp \;\&\; {\sim}p)$ 3, E.G.
1,2 (5) $\exists pF^1 p$ 4, F
1,2 (6) $\exists pF^n p$ 5, lemma
1 (7) $\exists pF^n p$ 2, 6, RAA
1 (8) $\exists pF^m P$ 7, lemma'

What does this result show, and why, indeed, should we be interested in obtaining it? Prior took his result, rightly I think, to show that there are strong and potentially unobvious self-reference restraints on what we can state, believe, think, hope, and so on. But, it might be argued, these restraints are easily avoided (indeed it has been so argued: see Christopher G. New, 'A note on the paradox of the preface', *Philosophical Quarterly* 28, October 1978, pp. 341–4). For while a writer cannot claim, in the context in which the *other* statements occur, that at least one of his/her statements is untrue, without running afoul of Prior's result, s/he may do so with impunity in *another* context. Thus someone may say, in a preface, 'at least one of the statements in the body of this work is false', or, in a later work, 'at least one of the statements in my earlier work was false'.

This is correct, but we might note in passing, first, that the result remains startling, and gives us an insight into propositional attitude terms, even if it can be avoided by careful writing, and, secondly, it is not at all clear what the equivalent of isolating contexts in this way would be like in the case of unuttered propositions. *Written* assertions can be isolated spatially, but what about held though unuttered beliefs? Of course, we can say, 'at least one of my *past* beliefs was false', but why should we debar ourselves from making the same claim about our *present* beliefs?

How might we isolate our fallibility admission from our *other* beliefs in such a case? There seem to be two possibilities. One is to try something like, 'at least one of my beliefs, other than this one, is false'. About this I share Ryle's namely-rider worries: it is far from clear that this sentence *does* record a belief. Indulging in a problem of self-reference is no way to avoid one. The other way is to *specify* the belief to be isolated: 'at least one of my beliefs is false, and the false belief is not the belief that at least one of my beliefs is false' (i.e., $D\,\exists p(Dp\ \&\ {\sim}p\ \&\ p \neq \exists p(Dp\ \&\ {\sim}p))$, which is to say, $D\,\exists pF^2p$). But, as we have seen, Prior's result can be obtained from this antecedent as well. And the generalized version of Prior shows that *no* attempt of this sort to block the result will succeed.

Just as there are logical restraints on what we can know, so that we believe ourselves to know things which we do not know, so there are logical restraints on what we can believe, so that we may believe ourselves to believe things which we do not and, in the circumstances, cannot believe. This gives us a further result, that some D operators, including *believes*, will not be suitable as analogues of possibility in S4-like systems, for there will be cases for which $DDp \rightarrow Dp$ does not hold.

This result also has consequences for scepticism. Consider the operator J, with the interpretation:

$J\alpha$: 'Conditions are such as to justify a belief by *a* that α'

As before, we have:

$$J\,\exists pF^np \rightarrow \exists pF^mp$$

i.e., if *a* is justified in believing that one of the (other) propositions *a* is justified in believing is false, then one of the (other) propositions *a* is justified in believing *is* false. Hence, since

$${\sim}\exists pF^mp \rightarrow {\sim}J\,\exists pF^np,$$

if it is *not* the case that one of the propositions *a* is justified in believing is false, *a* cannot be justified in believing that one of them is; again, a surprising limitation. In this area at least warrant must wait upon truth.

So far, then, we have seen, in the general case, that if we believe (hope, fear, assert, predict) in a given context that one of our beliefs (etc.) is false (unjustified, not going to be fulfilled), then:

(a) one of our beliefs (etc.) will be false (etc.).
(b) the false belief will not be *any* belief of the form $\exists pF^np$.

If we call such beliefs about the falsity of our beliefs meta-beliefs, then the false belief will not be any meta-belief.

(c) Indeed, it is a logical truth (line 10, lemma) that no such meta-belief is false.

(d) These results cannot be blocked by excluding any of the meta-beliefs from our antecedent, for $\exists pF^m p$ follows from any proposition of the form $D\,\exists pF^n p$.

If you *can* believe or state that there are such false propositions, then there are, and they're not meta-propositions. However, as a consideration of the contrapositive:

$$\sim \exists pF^m p \rightarrow \sim D\,\exists pF^n p \qquad \text{for any } n, m \geqslant 1$$

shows, the stating, or believing, or what have you, of such meta-propositions is harder, or might be harder, than most of us are inclined initially to believe, and it is at least an open possibility that some of the things we believe to be beliefs or propositions are not really so. Thinking that it's so in such cases may not really *make* it so, but at least when it's *not* so, we can't think that it is so, though of course we can think that that's what we think.

University of Calgary

© J. J. MacIntosh 1980

TRUTH AND KNOWABILITY

By J. L. MACKIE

W. D. HART, in 'The Epistemology of Abstract Objects', in *Aristotelian Society Supplementary Volume* LIII (1979), refers to an 'unjustly neglected logical gem' first published by F. B. Fitch in *Journal of Symbolic Logic* XXVIII. Hart claims that the argument in question can be used to disprove verificationism; at the Joint Session others claimed variously that it is fallacious and that it is not a proof but a paradox. However, none of these claims is quite correct, and the argument involves complications not shown in Hart's presentation.

To understand the argument, let us start with an operator-variable J, which has any number of innocent interpretations, such as 'It is written in green ink at t_1 that . . .'; by 'innocent' I mean that for pretty well any proposition p it is possible that Jp and also possible that $\sim Jp$. Then take K as the truth-entailing counterpart of J; that is, Kp is defined as Jp & p. Then it is tempting to say that for any p, it is possible that Jp, and hence that for any p which is true it is possible that Kp, and therefore to adopt the inference rule R, 'From p, infer $\Diamond Kp$'. Then provided that K distributes over conjunction, we have this proof:

1. $p \,\&\, \sim Kp$	supposition
2. $\Diamond K(p \,\&\, \sim Kp)$	from 1 by R
3. $\Diamond(Kp \,\&\, K(\sim Kp))$	from 2 by the proviso
4. $\Diamond(Kp \,\&\, \sim Kp)$	from 3 because K is truth-entailing
5. $\sim(p \,\&\, \sim Kp)$	from 1–4 by RAA, since 4 is false

This proof is plainly valid. It shows that there cannot be a truth-entailing operator K, which distributes over &, such that both $p \,\&\, \sim Kp$ and the inference rule R can hold. Where J is innocent, $\sim Jp$, and therefore also $\sim Kp$, will hold for many true propositions; there will be many truths of the form $p \,\&\, \sim Kp$. So if the proviso holds it is R that must be rejected. For example, we must reject the inference from p to 'It can be truly written in green ink at t_1 that p'.[1]

This is initially surprising. But only initially. It comes about because truth-entailing operators can be used to construct self-refuting expressions.[2] Thus while it may be true that

p, but it is not written in green ink at t_1 that p

---

[1] The generality of the argument was stressed by Jon Dorling, who also supplied the green ink example.

[2] In 'Self-Refutation: A Formal Analysis', in *The Philosophical Quarterly*, Vol. 14 (1964), pp. 193–203, I showed how various types of operator give rise to several different varieties of self-refutation.

it does not follow from this that

> It can be truly written in green ink at t_1 that (p, but it is not written in green ink at t_1 that p).

This should be no more surprising than the fact that while I may be saying nothing at t_1, I cannot say truly at t_1 that I am saying nothing at t_1.

Thus each of the interpretations of R that we get from innocent interpretations of J and hence of K is an unsound inference rule. It is simply not the case that whatever is true can be truly written in green ink at t_1, and so on. We may think that if whatever is true can be written in green ink at t_1, it follows that whatever is true can be truly written in green ink at t_1. But this does not follow, because there may be things that *are* true, and *can* be written in green ink at t_1, but which if they *were* written in green ink at t_1 would not be true.

Now one possible interpretation among others of K is 'It is known at some time that . . .'. Since this is truth-entailing and distributes over &, the proof will hold for it; there are surely many truths that are never known, so there must be truths of the form p & $\sim Kp$; so the R in question must be rejected. That is, we cannot infer 'It can be known at some time that p' from p; we must reject the principle that what is true can be known.

But does this disprove verificationism? Hart derives 'What is true can be known' from the three premisses 'What is true is meaningful', 'What is meaningful is verifiable', and 'What is verifiable can be known'; he accepts the first and third premisses, and so takes the rejection (by the previous argument) of 'What is true can be known' to require the rejection of 'What is meaningful is verifiable'. But (as R. G. Swinburne said in the Joint Session discussion) this way of using the argument refutes only a very strong form of verificationism in which 'verified' entails 'true'; otherwise 'What is verifiable can be known' does not hold.

However, this is not the end of the matter. Suppose that we read K as 'It is justifiably believed at t_1 that . . .' This will distribute over &, but we might expect the argument now to fail at step 4, since this K is not truth-entailing. But step 4 still goes through. If it is justifiably believed at t_1 that p is not justifiably believed at t_1, then p is *not* justifiably believed at t_1. So we can still use the argument to reject a proposed principle that what is true can be justifiably believed at t_1. On the other hand, if we read K as 'It is justifiably believed at some time that . . .', then step 4 does not go through. It does not follow that if it is justifiably believed at any time that p is not justifiably believed at any time, then p is not justifiably believed at any time. It might be justifiable at t_1 to think that p is false and never has been and never will be justifiably believed, and yet there might be some other time t_2 at which p was, or will be, justifiably believed. So the argument does not enable us to reject the principle that what is true can be justifiably believed at some time. So we cannot use such a conclusion to reject, by Hart's method, the principle that whatever is meaningful is verifiable in the sense that it can be justifiably believed at some time.

Alternatively, we might take K as 'It is true and verified at some time that . . .'. Then, if the sense of 'verified' is strong enough to ensure distribution over &, the argument now goes through, since this K is truth-entailing: it shows that we must reject the principle 'What is true can be both true and verified at some time'. If this were strictly analogous to such other applications of the argument as the green ink one, we should have to say merely that there may be things that *are* true, and *can* be verified, but which if they *were* verified would not be true—which would not refute verificationism. But it is worse (or, for an unreconstructed realist, better) than this. In fact a proposition of the form 'p but it is never verified that p' not merely cannot be true and verified, but simply cannot be verified (still given distributivity).[3] To verify a conjunctive proposition one would have to verify both conjuncts together. There would have to be some one time, say t_1, at which it was verified that p and it was verified that it is never verified that p, and this, as we saw in the similar case of justifiable belief, is not possible, despite the fact that 'It is verified that . . .' is not in general truth-entailing.

Consequently, if it is granted that something of the form 'p but it is never verified that p' is meaningful, we can immediately reject the principle that whatever is meaningful can be verified. But, as the disanalogy with the green ink example shows, this result comes not from the original argument, but from a slightly different use of some of its materials.

University College, Oxford © J. L. MACKIE 1980

---

[3] Michael Lockwood pointed this out, commenting on an earlier version of this paper.

'IF' AND 'EVEN IF'

By JOHN A. BARKER

I

IN ' 'Even if" ' (ANALYSIS 39.1, January 1979) Allen Hazen and Michael Slote argue in favour of the thesis that ' "if" and "even if" conditionals of the same kind have the *same* truth conditions, and that the function of "even" is to comment on some *pragmatic* feature of the context' (p. 35). As they point out (Footnote 2), some theorists, including John Pollock ('Four Kinds of Conditionals' *American Philosophical Quarterly* 12, 1975) have advanced the contrary view that 'even' affects the truth conditions of conditionals. In particular Pollock, among others, maintains that a subjunctive 'even if' conditional can be true only if its consequent is true. I will try to show that Hazen and Slote have failed to come to terms with the evidence for this opposing view.

II

In discussing the examples

(1) If you drank a gallon of whiskey, you'd get drunk
(2) Even if I (only) drank a quart of whiskey, I'd get drunk

the authors state: 'Since the characters in our whiskey example could easily be abstemious, it should be clear that 'even' can be used in true counterfactuals with false consequents' (p. 37). Pollock, however, has argued that such examples do not discredit his analysis:

> There are certain examples which appear to be counter-examples to this analysis. For example, we might say of a person "He would be fired even if he drank just a little" without meaning to imply that he will be fired. This appears to be an example in which "QEP" ["Q would be true even if P were true"] does not entail Q. However, I do not think that this is a real counter-example. "He would be fired even if he drank just a little" is a shortened form of "If he drank, he would be fired even if he drank just a little". The latter can be symbolized as "D > (FEL)" ["If D were true, then F would be true even if L were true"] (p. 53).

Applying the suggested treatment to (2), we obtain

(3) If I drank a quart of whiskey, I'd get drunk even if I did not drink more than a quart.

This translation is quite plausible. And even though the embedded 'even if' conditional has a false consequent, the 'even if' conditional is itself the false consequent of a true 'if' conditional. Hence the view that a subjunctive 'even if' conditional can be true only if its consequent is true remains unaffected.

Perhaps Hazen and Slote would resist parsing (2) as (3). Such resistance, however, could prove to be embarrassing. It is intuitively clear that given (3) and given

(4) I will not get drunk

we can *deduce*

(5) I will not drink a quart of whiskey.

Pollock can easily account for the validity of this inference, as he maintains that the consequent of (3) cannot be true if (4) is true. But since Hazen and Slote deny this, they must find some other way to justify the inference. The most promising strategy for them is to argue that since (3) is equivalent to, or at least entails (2), and since (5) follows from (2) and (4) by *modus tollens*, the validity of the deduction is established.

There is reason to believe, however, that even this strategy will not be successful. By *modus tollens* (2) and (4) entail

(6) It will not be the case that I (only) drank a quart of whiskey,

which seems to be equivalent to

(7) Either I will not drink a quart of whiskey, or I will drink more than a quart.

Clearly (7) does not entail (5). It is by no means certain, therefore, that even by appealing to the equivalence of (2) and (3), the authors can account for the validity of the deduction.

The case for Pollock's view becomes even more convincing when subjunctive 'even if' conditionals with *explicitly* compound antecedents are considered. For example,

(8) Even if I drank a quart of whisky and a gallon of coffee, I'd get drunk

is intuitively equivalent to

(9) If I drank a quart of whiskey, then I'd get drunk even if I drank a gallon of coffee.

From (9) and (4) we can deduce (5). But even though from an intuitive point of view (5) follows from (8) and (4), from a purely formal viewpoint the most that follows is

(10) Either I will not drink a quart of whisky or I will not drink a gallon of coffee.

If we accept Pollock's view that (8) is equivalent to (9) and that (4) entails the falsity of (9)'s consequent, we can justify the inference from (9) and (4) to (5), and from (8) and (4) to (5).

III

Although Pollock did not discuss *indicative* 'even if' conditionals, we can, I will argue, extend his view to cover indicatives. But I will approach the task indirectly, by way of examining a claim Hazen and Slote make about indicative conditionals.

The authors point out (p. 37) that an indicative conditional of the form

(a) If P, and if Q, then R

is ambiguous, as it can be read either as

(b) (If P then R) and (if Q then R)

or as

(c) If (P and Q) then R.

They go on to claim that

(d) If P, and even if Q, then R

can be read only as

(e) (If P then R) and (Even if Q, R).

But consider

(11) If I drink a quart of whiskey, and even if I drink a gallon of coffee, I will get drunk.

If (11) entailed the corresponding (e)-form sentence, it would entail the corresponding sentence of the form

(f) Even if Q, R.

Thus (11) would entail

(12) Even if I drink a gallon of coffee, I will get drunk.

Intuitively, (12) entails

(13) I will get drunk.

Evidently, an indicative 'even if' conditional, like its subjunctive counterpart, can be true only if its consequent is true. Since (11) does not entail (13), a (d)-form conditional should not be construed as being equivalent to the corresponding (e)-form conditional.

A (d)-form conditional, however, is equivalent to the corresponding conditional of the form

(g) If P, then even if Q, R,

and entails the corresponding conditional of the form

(h) If P then R.

For example, (11) is equivalent to

(14) If I drink a quart of whiskey, then even if I drink a gallon of coffee, I will get drunk,

and entails

(15) If I drink a quart of whiskey, I will get drunk.

An indicative conditional of the form

(i) Even if P, and even if Q, R

entails, but is not entailed by, the corresponding conditionals of the forms (d) and (g), as well as the corresponding conjunction of conditionals of the form

(j) (Even if P, R) and (Even if Q, R)

For example,

(16) Even if the bars are closed, and even if my liquor cabinet is empty, I will get drunk tonight

entails, but is not entailed by each of the following:

(17) If the bars are closed, and even if my liquor cabinet is empty, I will get drunk tonight

(18) If the bars are closed, I will get drunk tonight, even if my liquor cabinet is empty

(19) Even if the bars are closed, I will get drunk tonight; and even if my liquor cabinet is empty, I will get drunk tonight.

As in the case of subjunctive conditionals, an indicative conditional of the form

(k) Even if P and Q, then R

is ambiguous, as it can be read as

(i) Even if P, and even if Q, R

or as

(g) If P, then even if Q, R

or as

(h) If Q, then even if P, R

To illustrate, let us consider the three conditionals:

(20) Even if the bars are closed and my liquor cabinet is empty I will get drunk tonight
(21) Even if I drink a quart of whiskey and a gallon of coffee, I will get drunk
(22) Even if I drink a gallon of coffee and a quart of whiskey, I will get drunk.

Whereas (20) is to be interpreted as

(16) Even if the bars are closed, and even if my liquor cabinet is empty, I will get drunk tonight

(21) and (22) are to be interpreted as

(14) If I drink a quart of whiskey, then even if I drink a gallon of coffee, I will get drunk.

IV

In the case of 'even if' conditionals with negative consequents, the task of discerning the conditional's form can be more difficult. Consider the sentences

(23) If I (only) drink an ounce of whiskey, I won't get drunk
(24) Even if I drink two ounces, I won't get drunk
(25) I will get drunk
(26) I will drink more than two ounces.

Assuming that I will not drink any other intoxicating beverage besides whiskey, (26) follows from (24) and (25). It is clear that the antecedent of (24) is implicitly compound, and hence that (24) is equivalent to

(27) Even if I drink two ounces and do not drink more than two, I won't get drunk.

But (27) does not seem equivalent to

(28) If I drink two ounces, then even if I do not drink more than two, I won't get drunk,

which is a distinctly odd sentence, to say the least.

Evidently, the proper interpretation of (27), and hence of (24), is

(29) If I do not drink more than two ounces, then even if I drink two ounces, I won't get drunk.

This is not only an intuitively sound translation of (24), but clearly justifies the inference from (24) and (25) to (26).

V

I do not wish to suggest that Pollock's analyses of subjunctive 'even if' conditionals is totally acceptable. Nor do I wish to suggest that a satisfactory analysis of indicative 'even if' conditionals exists. Nevertheless, I think that there is strong if not conclusive evidence that 'even' affects the truth conditions of both subjunctive and indicative conditionals.

Southern Illinois University at Edwardsville © John A. Barker 1980

COMPATIBILISM AND THE BURDEN OF PROOF

By Peter van Inwagen

RICHARD Foley's 'Compatibilism and Control Over the Past' (Analysis 39.2) is an attack on James W. Lamb's 'On a Proof of Incompatibilism' (*The Philosophical Review* LXXXVI, 1977, pp. 20–35) and, rather less directly, on my 'The Incompatibility of Free Will and Determinism' (*Philosophical Studies* 27, 1975, pp. 185–199). In this note I shall try to bring out the oddness of a certain argument that Foley employs.

Foley's argument concerns a premise of Lamb's that Lamb calls 'a principle of can entailment':

> If a set F of true propositions logically entails that S does action A then if S at time t can refrain from A there is some member of F such that at t S can make it false.

In the course of a discussion of this principle, Foley says I present an argument in 'its' defence. Well, perhaps the argument of mine that Foley misquotes (p. 70) could be adapted to defend Lamb's principle, though my argument was for a principle of my own; what the relation is between my principle and Lamb's is a rather tricky question I shall not attempt to answer. (I do find Lamb's principle—let's call it 'PCE'—highly plausible whether or not it's mine, but I shall not defend it in this paper.)

Now for Foley's odd argument. He first presents a certain analysis of 'could have done otherwise' statements, a version of the well-known "conditional" analysis of such statements, and shows that if this analysis is right, PCE is wrong. He then argues as follows:

> Of course, I do not endorse the above analysis of 'could have done otherwise'. I have used it only to show how compatibilistic analysis of 'can' and 'could' statements might be used to argue against the principle of can entailment and accordingly show why Lamb and van Inwagen, if they wish their arguments to be convincing, must demonstrate the inadequacy of such analyses. But this neither Lamb nor van Inwagen has done. (Morover, *if* plausible arguments against compatibilistic analyses of 'can' and 'could' statements were available, then arguments using the principle of can entailment would seem to be unneeded. And so, it would seem that Lamb's and van Inwagen's argument *at best* is superfluous.) (pp. 71–72)

I shall first consider the parenthesis. I am not sure what Foley means by 'compatibilistic analysis of "can" '. Perhaps he means 'analyses that treat "can" statements as disguised conditionals'. In that case, he seems to be saying that if someone had a plausible argument for

> (1) All analyses that treat 'can' statements as disguised conditionals are wrong,

he would not need any further argument for

> (2) Free will and determinism are incompatible.

But at least one philosopher has accepted (1) and rejected (2) (see Keith Lehrer, 'An Empirical Disproof of Determinism?', in *Freedom and Determinism,* ed. K. Lehrer, Random House 1966). And whether or not that philosopher's reasons for accepting (1) and rejecting (2) were good, surely (2) does not follow from (1). A philosopher who wished to show that free will and determinism were incompatible could not accomplish this task by showing that some analysis, or class of analyses, of 'can' was wrong. Such an accomplishment could show no more than that a certain argument for compatibilism was unsound; it could not show that compatibilism was false. So much for the charge that Lamb's and my arguments are 'at best' superfluous. I turn now to Foley's contention that Lamb and I, if we wish our arguments to be 'convincing', must demonstrate the inadequacy of 'compatibilistic' analyses of 'can'.

Why must we do this? In order to accomplish *what*? It is not clear to me what general principles of argumentation Foley is appealing to. Does he accept this principle:

> No argument is convincing unless it is accompanied by arguments showing that all its possible counter-analyses are incorrect?

(Let us say that an analysis of a term or concept is a *counter-analysis* to an argument if it follows from the assumption that that analysis is correct that that argument is unsound.) If so, he is being unreasonable, since this principle obviously has the consequence that no argument is convincing, for no one *knows* all the possible analyses of any term. A more likely principle is this:

> No argument is convincing unless it is accompanied by arguments showing that all its *well-known* counter-analyses are incorrect.

More likely, perhaps, but I don't believe it. My argument and Lamb's appear to me to be perfectly convincing, though they are accompanied by no refutation of Foley's analysis or of any of the well-known family of conditional analyses of 'can'. It's true, of course, that my conviction that the premises of my argument are true depends on my *pre-analytical* understanding of 'can'. But then so does Foley's conviction that his analysis of 'can' is correct (or even worth considering) depend upon *his* pre-analytical understanding of this word. Suppose I were to say that Foley has to show that one of the premises of my argument was false if he wanted his analysis of 'can' to be convincing. Wouldn't that be a pretty odd thing for me to say? Isn't what Foley says *I* must do if I want my argument to be convincing odd for just the same reason?

I don't wish to appear intransigent, however. If Foley really must have an argument for the conclusion that his analysis is wrong, I'll give him one:

> PCE is true
> So Foley's analysis is wrong.

This strikes me as a pretty good argument. Foley himself admits that its conclusion follows from its premise. The premise, as Lamb says, verges on being tautological. And the conclusion is not particularly paradoxical or counter-intuitive. Nor, indeed, is the stronger statement (1), which I suppose also follows from PCE, particularly paradoxical or counter-intuitive. At any rate, *I* see no particular reason to think that 'can' statements are disguised conditionals. Does anyone?[1]

Syracuse University © PETER VAN INWAGEN 1980

---

[1] In the remainder of his paper, Foley does what he ought to have done in the first place and argues against PCE. I shall not discuss his argument. PCE is, after all, Lamb's principle and I will leave it to Lamb to defend it. In my papers 'Reply to Narveson' and 'Reply to Gallois', I have replied to certain criticisms of my own premises that are not entirely dissimilar to Foley's criticism of PCE. Both these papers appeared in *Philosophical Studies* 32 (1977).

REPLY TO VAN INWAGEN

By RICHARD FOLEY

IN 'Compatibilism and Control over the Past' (ANALYSIS 39.2, March 1979, pp. 70–4) I attack arguments against compatibilism forwarded by James Lamb and Peter van Inwagen. Van Inwagen has replied in 'Compatibilism and the Burden of Proof'.[1] His reply, however, misinterprets my attack.

Before showing how he misinterprets my position, a peripheral issue needs to be settled. Van Inwagen says that I misquote an argument of his. In fact I don't. The argument I quote is verbatim from his article. I did say that his argument is intended to support what Lamb calls 'the principle of can entailment'. Van Inwagen instead uses his argument to support a principle which he says is an instance of the following general principle (van Inwagen, *op. cit.* p 192):

> P: If *S* can render R false and if *Q* entails R, then *S* can render *Q* false.

Although *P* may differ in some subtle way from Lamb's principle, I can see no reason to think (and van Inwagen certainly has given me no reason to think) that the difference would in any way affect my attack.

So much for the peripheral issue. How has van Inwagen misinterpreted my position? Throughout most of his reply, he seems to think that I am offering an argument in favour of compatibilism. I do no such thing. My purpose is entirely negative: to show that van Inwagen has not resolved the compatibilist-incompatibilist controversy and that he has not even forwarded the controversy in a way which will help promote resolution. The problem is that van Inwagen uses the above principle P to construct a *reductio* against compatibilism. He uses P to argue that compatibilists are committed to the position that a person acts freely only if he is able to change the past or alter the laws of nature (since if determinism is true, past conditions and the laws of nature entail that a person acts as he does). Unfortunately, those not already convinced of the falsity of compatibilism can reasonably believe that there are counterexamples to P.[2] In particular, I pointed out that most compatibilists think they have good reasons to believe that (1) a person acts freely if he could have done otherwise and (2) 'could have done otherwise' can be analysed in such a way that a person could have done otherwise even if determinism is true. I then pointed out that given (1) and (2) compatibilists have no trouble generating counterexamples to van Inwagen's P. My purpose in pointing this out is not, as van Inwagen suggests, to advocate the absurd position that incompatibilists have an obligation to show that all possible 'counter-analyses' to their position are incorrect. It is simply to make explicit that compatibilists have a valid argument against a crucial premise in van Inwagen's argument and that unless there is some reason to think that seemingly plausible, commonly defended compatibilistic positions are inadequate, there is some reason to think this valid argument is also a sound argument.

---

[1] ANALYSIS 40.2, March 1980 pp. 98–100 Van Inwagen's original argument against compatibilism is in 'The Incompatibility of Free Will and Determinism', *Philosophical Studies* 27 (1975), pp. 185–199. Lamb's argument is in 'On a Proof of Incompatibilism', *The Philosophical Review* LXXXV (1977), pp. 20–35.

[2] I also argue that even if compatibilism is assumed to be false, there are counterexamples to P. See 'Compatibilism and Control over the Past', pp. 72–74.

Van Inwagen occasionally seems to realize that this is the point of my paper, but when he does he is surprisingly unconcerned. Why? Apparently he thinks it is, to use his word, odd for me to ask him to show that any compatibilistic analysis of 'could have done otherwise' is wrong. Indeed, van Inwagen also admits that it would be odd to ask a compatibilist to show that one of the premises of his (van Inwagen's) argument is false in order to show that a particular analysis of 'could have done otherwise' is correct. I confess that if we are going to argue by an appeal to oddness, it seems especially odd to me that van Inwagen would admit this. For he in effect is admitting that he has no argument which either is convincing or ought to be convincing to a compatibilist. All he can say is that P and the other principles he uses to argue against compatibilism seem analytic to him; he admits he has no way to argue for these intuitions which is 'both clear and non-question-begging' (see van Inwagen, 'Reply to Narveson', *Philosophical Studies* 32, 1977, p. 96). On this last point, van Inwagen and I agree.

So, van Inwagen gives us nothing other than an appeal to intuition as reason for thinking that P is true. Now, there clearly is a place for appeals to intuition in philosophical argument; I do not want to deny that. But I do want to claim that if we are interested in philosophical dialogue and the philosophical progress that is likely to accompany such dialogue, we should not look for premises which clearly are inconsistent with our opponent's position and then argue by appeal to intuition that because these premises are obviously true our opponent's position must be false. To do this is to encourage stalemate: each party in the dispute will assert dogmatically the falsity of the other's position. And this is precisely how van Inwagen defends P. He defends P, in effect, by three times calling my argument against it 'odd'. I am sure that what seems odd or true or analytic to van Inwagen is of the greatest importance to van Inwagen, but I cannot understand why it should be of much interset to compatibilists. P and other similar principles may seem analytically true to van Inwagen, but they seem patently false to compatibilists, and for just the reasons I mentioned.

The point here can be reinforced by contrasting van Inwagen's attack on compatibilism with those forwarded by J. L. Austin and Roderick Chisholm.[3] Both of these philosophers criticize compatibilists on their own ground, claiming that a certain class of commonly defended compatibilist strategies inevitably fail. And even if one does not think that these criticisms are ultimately successful, it is undeniable that they have had a beneficial effect on the compatibilist-incompatibilist debate. These criticisms have forced compatibilists to recognize that earlier compatibilist strategies were not very convincing and that better analyses must be forthcoming. In Chisholm and Austin we have a kind of criticism which the compatibilist is forced to take seriously. In van Inwagen we do not.

[3] J. L. Austin, 'Ifs and Cans', in *Philosophical Papers*, eds. J. O. Urmson and G. J. Warnock (Oxford: Clarendon Press, 1961); Roderick Chisholm, 'J. L. Austin's Philosophical Papers', *Mind* LXXIII (1964), pp. 20–25.

University of Notre Dame

© RICHARD FOLEY 1980

BERNARD WILLIAMS AND THE CARTESIAN CIRCLE

By A. C. STUBBS

IN Chapter 7 of his recent book on Descartes (*Descartes: the project of pure enquiry*, Penguin 1978) Professor Bernard Williams joins the ranks of those philosophers who have contended that the reasoning contained in the 'Third Meditation' is innocent of the charge of circularity commonly levelled against it. He draws, in his argument for this conclusion, upon the interpretation of Descartes' 'project of pure enquiry' which he expounds with characteristic depth of insight in Chapter 2 of his book. I shall argue that Williams fails to provide a successful defence of Descartes' procedure in the 'Third Meditation', and shall further suggest that it is this failure, not his supposed success, which is rooted in Williams' own account of the general aims of the *Meditations.*

The 'Cartesian Circle' is too familiar to require lengthy elucidation here. In brief, Descartes attempts to prove the existence of a God incapable of deceit in order to guarantee the truth of clearly and distinctly perceived propositions and thus vindicate his method of acquiring knowledge. However, since propositions of the type in need of divine guarantee figure as premises in Descartes' proof(s), his reasoning is circular, in that it presupposes what it sets out to prove—viz. that whatever is clearly and distinctly perceived is true. (Strictly speaking, this is a *consequence* of the conclusion of Descartes' arguments, not the conclusion itself, which is that a good God exists.) Some commentators have claimed to find in Descartes' own writings an answer to this 'circularity' objection, which Williams calls 'the memory answer' (p. 191), and, following H. G. Frankfurt (*Demons, dreamers and madmen: the defense of reason in Descartes's Meditations*, Bobbs-Merrill 1970, 'Memory and the Cartesian circle', *Philosophical Review* LXXI 1962), rightly rejects. According to this answer, the truth of what is clearly and distinctly perceived is never called into question by Descartes, who doubts only the reliability of his *memory* that he did, on some past occasion, clearly and distinctly perceive something. Williams cites ample and convincing textual evidence that, for Descartes, metaphysical doubt does extend (via the 'evil genius hypothesis') to the truth of what is clearly and distinctly perceived, and is only laid to rest by the proofs that a non-deceiving God exists. His own position does have one thing in common with 'the memory answer', however, viz. this: that no clear and distinct perception can be doubted *at the time of its occurrence*, i.e. one could not clearly and distinctly perceive that x and *at the same time* doubt that x. However, when not 'attending' sufficiently to x, when not 'thinking clearly' about what it involves, one can doubt it while *knowing* that one previously intuited its truth clearly and distinctly (pp. 186–187). This point is important, and should be looked at more closely.

Williams' view is that, for Descartes, some clearly and distinctly perceived propositions are 'irresistible', i.e. they 'have the property that if someone thinks of such a proposition, he believes it' (p. 185). It is in terms of this notion of *irresistibility* that Williams expresses the thought contained in the well-known passage from the *Second Replies* where Descartes writes of propositions that are 'so evident and at the same time so simple that we can never think of them without believing them to be true' (Williams' translation). Among these propositions are the *cogito* and the premises employed in the proofs of God's existence. It is important to note that both Williams' definition of 'irresistibility' and the remarks by Descartes on which it is based are couched in purely psychological terms; i.e. in terms having to do solely with the *attitude* of a thinker to certain propositions to which he is, at that moment, attending, and carry no implication that the propositions themselves are immune from metaphysical doubt concerning their truth. Williams goes on to say quite explicitly that 'Irresistibility does not entail truth' and that it may be 'a matter of psychological compulsion, one which the malicious demon would have been happy to implant' (p. 187). A proposition's being irresistible, then, comes down to this, that even though it may *not* be true, this is a possibility one can only entertain while not attending closely to it. We are dealing then with *psychological* irresistibility only, propositions which are merely *subjectively* certain; for *objective* certainty would entail truth. A proposition I find irresistible in this sense may yet be false, and this is so even if, while attending to it, there is no psychological room for the *thought* that it is so. I shall return to this crucial point later.

It may now look as if Williams must accept that there is, after all, a circle in Descartes' reasoning. Only a proof that a good God exists could, by eliminating the possibility of an omnipotent deceiver, transform psychological irresistibility into objective certainty; and how can Descartes construct such a proof without illicitly treating merely irresistible propositions as though they were objectively certain? However, Williams offers a line of argument which, he maintains, contains no such circle. He contends that Descartes, the pure enquirer, is committed by the nature of his project—the search for knowledge—to advancing from the 'time-bound' clear and distinct perceptions he already has by admitting 'some acceptance-rule for beliefs which are on-going and not time bound' (p. 201), viz. '(A) Accept as on-going beliefs just those propositions which are at any time clearly and distinctly perceived to be true' (p. 203). Descartes can then accept:

(1) God exists and is no deceiver

which entails:

(2) What one clearly and distinctly perceives to be true is true.

In consequence:

(3) (A) as an acceptance-rule for a truth-seeker is a sound rule.

While conceding that (2) and (3) are based on (1), Williams maintains that (1) is not likewise based on (2) and (3); it is arrived at 'merely by intuiting the reasons for it' (p. 203).[1] Hence there is no circle, *despite the fact that the intuition of* (1) *cannot validate it unless* (2) *and* (3) *are true*; for (2) and (3) do not have to be *known* before (1) can be known.

If I have correctly understood Williams, he sees Descartes as presenting the sceptic with two things:

---

[1] Williams' argument here bears some similarity to that of Anthony Kenny, as presented in his article 'The Cartesian Circle and the Eternal Truths', *Journal of Philosophy* 1970, and in Chapter 8 of his book *Descartes: A Study of his Philosophy*, Random House 1968. Kenny distinguishes between particular clear and distinct perceptions and the general rule that 'everything I clearly and distinctly perceive is true', maintaining that the object of metaphysical doubt is the general rule only. Thus there is no circle, as it is particular clear and distinct perceptions (which do not require divine guarantee) which are used in the proofs of God's existence. Williams is right as against Kenny in maintaining that particular clear and distinct perceptions *can* be doubted (though not at the moment of having them) but wrong, as I shall show, in thinking that he can use their 'irresistibility', plus (A), to escape from the Cartesian Circle.

1. A proof (or proofs) of God's existence which he cannot help but believe while attending to it (or them).
2. An acceptance-rule (A) for beliefs which, while 'a merely structural condition on having usable knowledge' (p. 206) blocks any attempt by the sceptic to re-introduce doubt while not attending to the proofs and thus not actually experiencing their irresistibility.

Systematic and general doubt, Williams claims, is thus pre-empted by a systematic and general answer in terms of the existence of a non-deceiving God (p. 207).

I shall now attempt to show that Williams' argument does not work. Let us return firstly to the notion of an irresistible proposition, on which so much of his argument hangs. We have seen that it must be psychological, subjective irresistibility that is meant here, i.e. irresistibility is simply a matter of what I cannot, in fact, help but believe. The *proposition* which I cannot help but believe may not be *true*, despite the fact that I cannot doubt its truth while actually thinking about it. Williams sees, of course, that 'irresistibility' alone will not provide us with an answer to the sceptic, but he locates its inadequacy in the fact that it is *time-bound*. Irresistible propositions are only irresistible for as long as I am thinking clearly about them. Hence the need for (A), which attempts to convert them into on-going beliefs. It seems to me, however, that the inadequacy lies in the *kind* of irresistibility clearly and distinctly perceived propositions have, not in the fact that they cease to be irresistible when I cease to attend to them. The trouble is that even when I am convinced of their truth, this is an entirely subjective matter—and what Descartes is after, and what the sceptic denies that he can get, is *objective* certainty. With this in mind, let us go back to the confrontation between Descartes and the sceptic as Williams presents it. Descartes is now offering the sceptic a proof of God's existence which cannot but convince him while he is attending to it. Suppose he ceases to attend to it, and observes that since irresistibility does not entail truth, 'a non-deceiving God exists' is not objectively certain, and that therefore it is still the case that, as Descartes says in the 'First Meditation', there may be 'some evil genius, not less powerful than deceitful [who] has employed his whole energies in deceiving me' (Haldane and Ross's translation). Descartes now has to do more than 'invite him to go through the proofs' again (p. 206); this really *would* be analogous to inviting him to sniff once more 'some hallucinogenic gas' (p. 207). He appeals to his acceptance-rule for beliefs (A). The sceptic has two answers to this. He can (1) point out, as I have suggested above, that acceptance of (A) would not remedy the *real* deficiency attaching to an irresistible proposition—it would simply convert a time bound subjective certainty into an on-going one; and he can also maintain (2) that, in any case, (A), as Williams concedes, is *itself* ultimately revealed as sound only by the very proof of God's existence now under attack by the sceptic as open to *metaphysical* doubt even at such times as one cannot (psychologically) *think* that it is ('if a proposition is irresistible it cannot be doubted' is of course an analytically true statement; but this does not mean that we are dealing here with *logical* irresistibility).

Although Williams does concede that (A) is only sound if (1) is true, he maintains that to reject its use is simply to refuse 'a merely structural condition on getting on at all' (p. 206). Waiving the point about how much (A) would actually achieve if accepted, why cannot the sceptic reply: 'precisely; in the face of the evil genius hypothesis we *cannot* get on at all. Knowledge is impossible'. Williams speaks of the sceptic as objecting '*merely* because he is no longer intuiting' (p. 206; my italics) and goes on: '. . . it would be unreasonable to spend all one's time rehearsing the proofs of the general answer to scepticism, an answer which we nevertheless possess.' But the point is that the sceptic denies with perfect consistency that we *do* possess the answer. He takes his stand on the merely *subjective* certainty possessed by the proofs which ultimately justify reliance on clear and distinct perceptions and hence the soundness of (A). The premises of these proofs cannot possess metaphysical certainty (as opposed to psychological irresistibility) until the conclusion is reached; thus we are right back with the circle. Thus the sceptic is not, as Williams seems to suggest, asking for continuous rehearsals of the proofs of God's existence; he is maintaining (rightly) that no proof, however continuously rehearsed, which is irresistible *only in a psychological sense* can constitute an 'objectively valid' refutation of the evil genius hypothesis. Without such a refutation, the possibility of knowledge has not been demonstrated. Williams is correct in saying that, as far as the possibility of systematic error in our representation of reality is concerned 'Descartes holds that here it makes a difference whether [a person] A believes in God or not' and that he is consistent in holding this. However, the sceptic is equally consistent in demanding more than *belief*, however subjectively irresistible. He requires objective certainty, certainty which can withstand even the most extreme, hyperbolical doubt.

It is at this point that it is appropriate to refer to Williams' explication of Descartes' 'project of pure enquiry' in Chapter 2. Descartes is correctly presented here as a seeker after knowledge, as involved in a quest for truth which 'seeks certainty against any *conceivable* doubt' . . . a method of enquiry 'which should be error-proof' (p. 66). This extreme requirement springs from the 'absolute conception' of reality which is sought by the pure enquirer; from this viewpoint 'the distinction between a source of error or distortion which is merely conceivable and one which we take to be empirically effective loses its importance'

(p. 67). Hence the Cartesian sceptic introduces into the 'First Meditation' a reason for doubt which Descartes refers to as 'slight' and 'metaphysical', but which has devastating consequences; viz. the bare possibility that our representations of reality are systematically distorted by an omnipotent and malicious being bent on deceit. When we turn from Chapter 2 to Chapter 7, we find Williams talking in terms more reminiscent of Frankfurt's (rejected) conclusion summarised by Williams on p. 199: 'the point is, not that a non-deceiving God exists, but that reason leads to this conclusion'. The persistent sceptic, Williams now says, is someone 'uninterestingly beyond the reach of rational discourse', someone who refuses to acknowledge 'the result of a simple and maximally clear argument' (p. 207). But is not the point that, in the face of the evil genius hypothesis, we cannot *trust* simple and maximally clear arguments, we cannot know that what we cannot help but believe is *true*? And it is *truth*, as Williams himself emphasises, which is the goal of Descartes' project of pure enquiry, truth defined in terms of correspondence with an external, absolute reality.

I do not wish to suggest that the Cartesian sceptic is in an attractive position, or that there is no more that we can say to him (or at least, to ourselves *about* him). We can argue along now familiar lines that the evil genius 'hypothesis' is not a hypothesis at all, since nothing could count for or against it. We can say that the radical sceptic's doubts are *idle*, that they make no difference to anything, and even that, in consequence (by a swift verificationist move) they are meaningless, unintelligible. These are now familiar strategies employed against scepticism. The difficulty with such strategies is that they tend to lead us away from something both Descartes and Williams would want to keep, viz. the 'absolute conception' of reality, the Archimedean point for which the pure enquirer ultimately seeks.

The Queen's University of Belfast

© A. C. STUBBS 1980

RIGHTS, GROUPS AND DISCRIMINATION: A REPLY TO KETCHUM

By ROBERT L. SIMON

IN a recent paper in ANALYSIS ('Statistical Justifications of Discrimination,' ANALYSIS 38.1, January 1978, pp. 37–42) I claimed that arguments like (a) below

> (a) 80% of the members of group X were victims of past discrimination. Therefore, for purposes of administrative convenience, all members of X should be regarded as eligible for compensatory benefits.

are logically parallel to arguments like (b).

> (b) 80% of women are physically incapable of performing task T. Therefore, for purposes of administrative convenience, no women ought to be considered as candidates for T.

I concluded that since (b) is unacceptable, and since (a) is logically parallel to (b), (a) is unacceptable as well. Accordingly, the so called administrative defence of preferential treatment, which presupposes the acceptability of arguments like (a), fails (see James Nickel, 'Should Reparations be to Individuals or Groups?' ANALYSIS 34.5, April 1974, pp. 154–160).

In reply, Sara Ann Ketchum claims that I am committed to a faulty analysis of race and sex discrimination and that I 'seem to assume without argument that the white males who would get jobs if there were no policy of preferential treatment have a right to those jobs and that those jobs would be so to speak "stolen" from them . . . on such a policy' ('Evidence, Statistics and Rights: A Reply to Simon,' ANALYSIS 39.3, June 1979, p. 150). Since I neither presented an analysis of race or sex discrimination nor advanced a view as to the rights of the most qualified job applicant, I fail to discern why Ketchum attributes these views to me or thinks they are relevant to my original argument.[1] However, as Ketchum's rejoinder does raise substantial questions about the issue I did address, the acceptability of inferences from premises about groups to conclusions about treatment of individuals, I believe her arguments are worth pursuing further.

---

[1] In fact, I find Ketchum's own analysis of systematic and pervasive discrimination quite attractive. And while I do think the most qualified applicant has a *prima facie* right to the job, I do not see that my original argument relies on such a claim. More important, I believe such a right can be overridden for a variety of reasons, including compensatory ones. Finally, in no sense do I believe beneficiaries of preferential treatment have stolen, or 'stolen', jobs from the most qualified applicants. For useful discussion of the rights of the most qualified job applicant, see chapter two of Alan H. Goldman's *Justice and Reverse Discrimination*, Princeton University Press, 1979.

Ketchum relies on such an inference at a particularly key point. For she argues that since white males as a group have acquired advantages from unjust discrimination against women and minorities, preferential treatment of particular women and minorities over particular white males now is justified. 'The suffering that Simon claims the young white males to be paying for the sins of their elders turns out to be the suffering of those who would not be able to benefit as much from past injustices as they had hoped or expected.'[2]

At least two questions need to be raised about this argument. First, is it true that *all* white males have secured *net* advantages from unjust discrimination against women and minorities? Second, if such a claim is true, does it follow that preferential treatment of women and minorities is justified as a means of nullifying the unfairly gained advantages of white males?

Surely, there are strong grounds for doubting the claim that all white males have secured net advantages from sex and race discrimination. For one thing, such systematic and pervasive forms of injustice may have contributed to a social climate in which other forms of injustice, some of which have victimized white males, have flourished. Consider, for example, the alleged connection between racial discrimination and economic exploitation of whites in the pre-Civil-War South. Moreover, as Alan Goldman has pointed out, discrimination entails inefficiency (the best candidates are not hired or admitted) and such inefficiency may have harmed some white males more than discrimination against others helped them (Goldman, *Justice and Reverse Discrimination*, p. 108–109). Still other white males may have to overcome insidious forms of discrimination themselves and may have claims to redress of their own. Indeed, some affluent white females may have benefited more from systems of injustice and exploitation than have some white males.

Accordingly the claim that all white males have secured net advantages from past discrimination (or that they have in all cases secured more of an advantage than any woman) is open to serious objection. Thus unless Ketchum means to advance arguments like (a) and (b), and argue from statistical correlation, her argument appears to rest on a doubtful premise. Even if *most* white males have (however inadvertently) secured net advantages from past discrimination against women and minorities, it does not follow that *every* white male may legitimately be disadvantaged by policies of preferential treatment. For the white male disadvantaged by such a policy may not be among the subclass of white males who have profited from the victimization of others.

---

[2] Surely the most Ketchum is entitled to claim is that white males expect and hope to profit from a system that in fact operates unjustly. She is not entitled to claim, at least not without providing some supporting evidence, what her language suggests: namely, the entirely different claim that white males intend and hope to profit from what they perceive as injustice.

Suppose, however, that in fact *all* white males have secured net advantages from past unjust discrimination. It hardly *follows* that the kinds of preferential treatment by race and sex now at issue in hiring or graduate school admission are justified. For one thing, such practices impose the costs of redress arbitrarily. If all white males are liable, it is hardly fair that a small proportion pay the entire cost of restitution while the overwhelming majority give up nothing. More important, even if all white males have benefited, they have not all benefited *equally*. Yet preferential treatment imposes costs according to one's relative lack of qualifications in the market, not degree of benefit received.[3] Indeed, those adversely affected are those with relatively weak records, precisely those who would tend to have benefited least from past injustice or have been victims of it themselves.[4]

What this suggests is that generalizations about groups do not easily lend themselves to direct support of conclusions about the just treatment of individuals. If we are to treat similar cases similarly, we must be sensitive to differences as well as the resemblances between group members. At the very least, at this point in the debate over preferential treatment much more argument than Ketchum provides is needed if her position is to avoid the kinds of objections sketched above.

At this point, however, it may be objected that I have unjustifiably assumed an individualist framework of discussion. For isn't it Ketchum's point that discrimination of a systematic and pervasive kind affects individuals in virtue of their *group* membership? If so, the very point of so called 'reverse discrimination' is not to deal out fair redress on an individual basis but to restore to victimized groups the social and economic position they would have held in the absence of systematic injustice.

If this is Ketchum's point, it seems to me to parallel Paul Taylor's recent defence of group compensation. ('Reverse Discrimination and Compensatory Justice,' ANALYSIS 33.6, June 1973, pp. 177–182.) I will not add to the criticisms of that view that have been advanced elsewhere except to suggest that even if our goal is compensation of groups we must still be as fair as possible to individuals in the process. Thus, Taylor's original proposal calls for *society* to compensate victimized groups. I have argued above, however, that the kind of preferential treatment currently at issue imposes costs arbitrarily on *individuals*. Accordingly, an appeal to group considerations along lines advanced by Taylor will not support Ketchum's position. It is unfair that arbitrarily selected individuals pay the entire cost of promoting equality for groups.

---

[3] I have made a similar point about the beneficiaries of preferential treatment in my 'Preferential Hiring: A Reply to Judith Jarvis Thomson,' *Philosophy & Public Affairs*, 3 : 3 (1974) p. 315. This point also counts against the acceptability of Ketchum's inference from the claim that every member of certain victimized groups have been injured by systematic discrimination to the acceptability of a system of preferential treatment which awards benefits according to qualifications, not degree of injury suffered.

[4] Goldman makes a similar point about the beneficiaries of preferential treatment, in his *Justice and Reverse Discrimination*, pp. 90–91.

I have suggested, then, that excluding particular white males from consideration for various benefits simply because white males as a group have benefited from past injustice raises serious issues about stereotyping and insensitivity to individual circumstances. Appeal to statistical correlations will not help us avoid these issues for the reasons given in my earlier paper. Accordingly I conclude that, for all Ketchum has said, present recommendations for preferential treatment by race or sex raise many of the same questions of injustice that properly apply to the invidious discrimination practised for so long against women and minorities.

Hamilton College © ROBERT L. SIMON 1980

New Paperbacks from Basil Blackwell

Philosophical Grammar

Ludwig Wittgenstein

Edited by Rush Rhees, translated by Anthony Kenny

Wittgenstein wrote the *Philosophical Grammar* during the years 1932–1934, the period just before he began to dictate the *Blue Book*. Although it is close to the *Investigations* in some points, and to the *Philosophische Bermerkungen* at others, the *Philosophical Grammar* is an independent work and discusses much that is not in either of them.
496 pages, paperback about £4.95

Logical Investigations

Gottlob Frege

Edited with a preface by P. T. Geach, translated by P. T. Geach and R. H. Stoothof

The three articles translated in this volume: 'Thoughts' (*Der Gedanke*), 'Negation' (*Die Verneinung*), and 'Compound Thoughts' (*Gedankengefüge*), were previously published by Frege in the periodical *Beiträge zur Philosophie des deutschen Idealismus,* the first two in 1918 and the last one in 1923. They were intended as chapters of a book, to be called *Logische Untersuchungen,* which Frege never finished.
92 pages, paperback £2.95

Form and Content

Bernard Harrison

'The book's interest resides in an original approach to the nature and problems of colour names, and in a novel application in this area of a Chomskyan model of linguistic competence.' David E. Cooper in *Mind*

'Well written and subtly argued, this book is a welcome addition to philosophical literature.' *Library Journal*

160 pages, paperback £2.95

Perception, Emotion and Action

A Component Approach

Irving Thalberg

In an attempt to achieve a clearer understanding of perception, emotion and action, the author begins by accepting at face value twelve lively but puzzling controversies about these phenomena. Each debate concerns the relationship between an event which seems to be in some mysterious way involved with perception, emotion or action. Quite apart from the sort of philosophical analysis they generate, the riddles treated here provide challenging intellectual exercises for anyone who is curious about the relationship between seeing and the brain processes which are indispensable to it.
150 pages, paperback £2.95

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, J. H. Benson; T. R. Baldwin, Margaret A. Boden, A. E. Fisher, Andrew Harrison, R. F. Holland, Hidé Ishiguro, J. Kemp, J. H. McDowell, Bernard Mayo, D. H. Mellor, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

SUBSCRIPTIONS. The subscription to ANALYSIS for institutions is £8.00 (inland), £9.80 (overseas), $24.00 (U.S.A.), $28.00 (Canada); for individuals £5.00 (inland), £6.00 (overseas), $15.00 (U.S.A.), $17.50 (Canada). Each volume comprises four numbers, three of 48 pages and one of 64 pages, appearing within the academic year—in January, March, June and October. Orders should be sent to Basil Blackwell, 108 Cowley Road, Oxford OX4 1JF, or placed with any bookseller.

CONTRIBUTIONS. **Articles submitted for publication should be addressed to Mr. Christopher Kirwan, Exeter College, Oxford OX1 3DP.** Contributors are asked to note the following requirements.

Articles should **normally** not exceed 3,000 words in length. **Occasionally,** however, longer contributions can be accepted;

They must be type written in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

It is regretted that owing to increased postal charges it is no longer possible to return typescripts unless the following instructions are followed:

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgement is required, a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope **and international reply coupons** of the requisite value, whether for air or surface mail.

Galley proofs of accepted articles will be sent to authors for correction, together with information about offprints. Typescripts will be retained by the Editor on the assumption that authors have kept their own copies.

The copyright of articles printed in ANALYSIS remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.

ISSN 0003-2638

PRINTED IN GREAT BRITAIN BY BURGESS & SON (ABINGDON) LTD., ABINGDON, OXFORDSHIRE

27 MAY 1983

Vol. 40 No. 3 (New Series No. 187) June 1980

ANALYSIS

Edited by
Christopher Kirwan

CONTENTS

Peacocke and Kraemer on Butler's problem — E. J. Lowe
Justice, truth and Nozick — Bernard Mayo
Counting proper classes — James Robert Brown
What is, necessarily is, when it is — C. J. F. Williams
Rationality and intransitive preferences — R. I. G. Hughes
Referential shifts — Rod Bertolet
Screening-off and the temporal asymmetry of explanation — A. David Kline
Materialism and the unity of consciousness — Andrew Ward
Counterfeit coins and forged paintings: caveat emptor — Michael Wreen
The ship of Theseus — Theodore Scaltsas
Frank Jackson and the spatial distribution of sense-data — A. Olding
A speech act dilemma — C. R. Carr
McHarry's theodicy: a reply — R. K. Perkins, Jr.
Lifesaving — T. M. Reed
You'll regret it — I. L. Humberstone

BASIL BLACKWELL · ALFRED STREET · OXFORD

The Philosopher's Index

Richard H. Lineback, *Editor*

The Philosopher's Index is an **up-to-date quarterly** index of articles from more than 300 international philosophy and interdisciplinary journals. Professional philosophers carefully index all articles, which are listed alphabetically by subject and author. Abstracts of the articles are published in each issue and are generally written by the authors of the articles. The Book Review Index contains an alphabetical listing by authors of recently published philosophy books, which have been reviewed in the journals indexed. A **cumulative edition** of *The Philosopher's Index* is published annually in the spring following the volume year. It contains in merged form alphabetized listings by subject and author of all the data published in the preceding four quarterly issues.

Special Editions

The Philosopher's Index: A Retrospective Index to U.S. Publications from 1940 includes approximately 15,000 articles from U.S. journals published during the 27 year period, 1940-1966, and approximately 5000 books published during the 37 year period, 1940-1976. This *Retrospective Index* was published in April 1978 with 1619 pages in three hardbound volumes.

The Philosopher's Index: A Retrospective Index to non-U.S. English Language Publications from 1940 includes approximately 12,000 articles published during 1940-1966 and approximately 5,000 philosophy books published during 1940-1978. This *Retrospective Index* is scheduled to be published in February or March 1980 with about 1400 pages in three hardbound volumes.

Their broad scope makes these special editions essential research tools in philosophy. Author and subject indexes are provided. The subject indexes provide an average of six subject headings per article and twelve per book. An abstract for each entry has been included, when available from the author or other sources. Recognizing the need for these publications, the National Endowment for the Humanities has supported the preparation of these retrospective editions by Grant RT-23984-76-375 and Grant RT-27265-77-1360.

Expanded Coverage

The Philosopher's Index will significantly expand coverage in 1980. In the past, current issues of *The Philosopher's Index* indexed journal articles only. However, Volume XIV (1980) will increase coverage to include original scholarly books in philosophy, plus bibliographies, biographies, translations, textbooks, dissertations, and anthologies. Each contribution to an anthology will be indexed as well as the work as a whole.

Subscription Information

The Quarterly, Paperback, $60 (Individuals: $30) / Annual Cumulative Edition, Hardbound, $70 (Individuals: $35) / *Retrospective Index to U.S. Publications*, Hardbound in 3 volumes, $195 (Individuals: $130) / *Retrospective Index to non-U.S. English Language Publications*, Hardbound in 2-3 volumes, $170 (Individuals: $113)

U.S. and Canada:

THE PHILOSOPHER'S INDEX
Bowling Green State University
Bowling Green, Ohio 43403 USA

Outside the U.S. and Canada: D. Reidel Publishing Company/Dordrecht, Holland

INQUIRY

Editor: Alastair Hannay
Founding and Consulting Editor: Arne Naess

Volume 22 (1979), No. 4

Are Luddites Confused?	Dan Lyons
The Centrality of Problem-Solving	John Kekes
The Roots of Epistemological 'Anarchy'	Nancy J. Nersessian
Sherlock Holmes – Philosopher Detective	Wulf Rehder
Making Sense of History: Skagestad on Popper and Collingwood	M. Hurup Nielsen and Jeremy Shearmur

Volume 23 (1980), No. 1

Democracy, Elitism, and Scientific Method	Paul Feyerabend
Science, Reason, Knowledge, and Wisdom: A Critique of Specialism	Nicholas Maxwell
Creative Product and Creative Process in Science and Art	Larry Briskman
Epistemological Relativism in its Latest Form	Harvey Siegel
Siegel on Kuhnian Relativism	Gerald Doppelt
A New Science of Criticism	Price Charlson

Subscription:

Foreign: Institutions: N.kr 130,–, $26.00 – Individuals N.kr 78,–, $16.00.
Norway: Institutions: N.kr 98,–, Individuals: N.kr 58,–.

Published quarterly by Universitetsforlaget, P.O. Box 2959 Tøyen, Oslo 6, Norway – P.O. Box 258, Irvington-on-Hudson, New York 10533, USA.

PHILOSOPHICAL BOOKS

Philosophical Books aims to provide prompt, scholarly reviews of professional books in philosophy. Besides conventional reviews in both contemporary philosophy and the history of philosophy, it now includes regular Discussion features, where the authors of selected titles have the opportunity to reply to reviewers. There are occasional subject surveys and issues devoted to a single theme.

Books for review should be sent to the Editor, Dr M.A. Stewart, Department of Philosophy, University of Lancaster, Lancaster LA1 4YT.

Annual subscriptions should be paid to the publishers:
Individuals £8 (overseas £9.50, US $20, Canada $23)
Institutions £12 (overseas £14.50, US $30, Canada $35).

There are now *four* issues a year.
Single copies and back issues are available.

BASIL BLACKWELL PUBLISHER
Journals Department, 108 Cowley Road, Oxford OX4 1JF

RATIO
edited by MARTIN HOLLIS

Volume XXII *June 1980* *No. 1*

CONTENTS

DAVID MILLER — *Justice and Property*
NANCY DAVIS — *Utilitarianism and Responsibility*
TERRANCE C. McCONNELL — *Utilitarianism and Supererogatory Acts*
DAVID CARR — *What Place has the Notion of a Basic Action in the Theory of Action?*
JOHN P. CLEAVE — *Some Remarks on the Interpretation of 3-Valued Logics*
JOSEPH AGASSI and CHARLES SAWYER — *Was Lakatos an Elitist?*
WILHELM BAUM — *Ludgwig Wittgenstein's World View*
YEHUDAH FREUNDLICH — *Who Cares About Identity Across Possible Worlds?*
LEROY N. MEYER — *A Note on Singularity and Exactness*
SUSAN JAMES — *Review of Karl Marx's Theory of History* by G.A. Cohen

Annual Subscription: £15.25 (inland) £18.75 (overseas) $41.25 (USA) $47.75 (Canada)
Orders to: Journals Department, Basil Blackwell Publisher Ltd.,
108 Cowley Road, Oxford OX4 1JF, England.

The German edition of *Ratio* is published by FELIX MEINER VERLAG,
Richardstrasse 47, 2000 Hamburg 76, West Germany.

Published by BASIL BLACKWELL, Oxford

Journal for the Theory of Social Behaviour

Edited by P.F. SECORD, (Department of Urban Studies, Queens College, CUNY) and R. HARRE (Linacre College, Oxford). Associate Editors, P. MANICAS (Department of Philosophy, Queens College, CUNY) and C. SMITH (Department of Sociology, Queens College, CUNY).

Volume 10, No. 1 March 1980

W.H. FODDY and W.R. FINIGHAN: The Concept of Privacy from a Symbolic Interaction Perspective 11
ARDA DENKEL: The Speaker's Communicative Intent 19
T.A. NOSANCHUCK: The Psychologic of Implication 39
V. REYNOLDS: *Review Essay:* Animal Behaviour and Human Nature 57

Annual subscription: Institutions £17.00 (inland) £20.75 (overseas)
$43.50 (USA) $50.50 (Canada); Individuals £9.00 (inland)
£11.00 (overseas) $22.50 (USA) $26.00 (Canada)

Orders to: JOURNALS DEPARTMENT, BASIL BLACKWELL & MOTT LTD., 108 Cowley Road, Oxford OX4 1JF, England.

PEACOCKE AND KRAEMER ON BUTLER'S PROBLEM

By E. J. Lowe

IN his recent book *Holistic Explanation* (Oxford University Press, 1979, p. 74), Christopher Peacocke suggests that R. J. Butler's ANALYSIS Problem No. 16 (ANALYSIS 37.3, March 1977) is the sort of problem which 'care in specifying the description under which an event is intentional . . . simply dissolves'. While not wishing to rehearse my own attempted solution to that problem ('Neither Intentional nor Unintentional', ANALYSIS 38.3, June 1978), I do want to take issue with Peacocke's suggestion.

The problem, it may be recalled, was this:

> If Brown in an ordinary game of dice hopes to throw a six and does so, we do not say that he threw a six intentionally. On the other hand if Brown puts one live cartridge into a six-chambered revolver, spins the chamber as he aims it at Smith and pulls the trigger hoping to kill Smith, we would say if he succeeded that he had killed Smith intentionally. How can this be so, since in both cases the probability of the desired result is the same?

Peacocke remarks: 'The alleged problem simply arises from not comparing like with like. The shooter has control over the conditional that if the gun fires, then it fires in one direction rather than another. But he does not have control over whether the gun fires at all. So we say that he intentionally shot Smith, but not that it was intentional of him that a full rather than an empty chamber was engaged. Now in the dice case, the agent has control over whether, for instance, the die falls on one side of the table with the six uppermost *if* it falls with the six uppermost at all, but he does not have control over the antecedent of this conditional. We say in this case exactly what we would expect to say by analogy with the gun example: it may be intentional of the agent that the die falls on the left-hand side of the table, but not that it falls with a six uppermost. Butler's problem arises only when we compare a condition over which the agent does have control in the gun case with a non-corresponding condition in the dice case over which he does not have control.'

Now it may not be immediately clear how this alleged solution (or dissolution) of the problem simply results from taking 'care in specifying the description under which an event is intentional', since this consideration nowhere explicitly enters Peacocke's discussion. (Peacocke's solution is, in fact, very similar both to Eric Russert

Kraemer's, 'Intentional Action, Chance and Control', ANALYSIS 38.3, June 1978 and to David Ross's, 'He Loads the Gun, not the Dice', ibid. Kraemer and Ross also make no explicit appeal to the relativity of intentionality to a description.)

Presumably, however, what Peacocke means to suggest is something like the following. In the dice game Brown's action may be characterized by various different descriptions, such as, let us say,

A1. 'throwing the die'
A2. 'throwing the die on the left-hand side of the table'
A3. 'throwing the die so that it falls with the six uppermost'
A4. 'throwing a six'
A5. 'winning the dice game'

(This list may be compared with that given by Kraemer who, however, makes no suggestion that his is a list of different descriptions of a single action, regarding it rather as a list of different 'states of affairs'.) Now Peacocke would presumably contend that Brown's action is intentional under descriptions A1, A2 and possibly A5 (Kraemer certainly has a presumption in favour of the latter) because with respect to these descriptions Brown has a requisite measure of 'control' (whatever that may exactly mean), whereas Brown's action is not intentional under descriptions A3 and A4 because with respect to these such control is absent.

Analogously in the gun example Brown's action may perhaps be characterized by the following list of descriptions (again comparable with a list of 'states of affairs' given by Kraemer):

B1. 'aiming the gun at Smith and pulling the trigger'
B2. 'engaging a full chamber and discharging it'
B3. 'firing the gun'
B4. 'shooting Smith'
B5. 'killing Smith'

This time Peacocke would presumably contend that Brown's action is intentional under descriptions B1, B4 and B5, but not intentional under descriptions B2 and B3 (and Kraemer would, in his own terms, clearly concur).

The 'dissolution' of Butler's problem might then be thought to boil down to this. There only appears to be a problem at all if we suppose (because we fail to realize that it is only meaningful to ascribe intentionality to an action *under a description*) that since in one of the cases Brown's action may be characterized as intentional (under a certain description) it is paradoxical that in the apparently parallel case it should not be so characterized (under a certain description). This, it may be urged, would only be paradoxical if it were the case, which it is not, that an action which may be characterized as intentional under one description could not also be characterized as not intentional under another description. For in fact in *both* cases Brown's action is intentional under some descriptions and not intentional under others.

If this is the sort of 'dissolution' Peacocke has in mind, then it is clear that it is inadequate and that the solution which he explicitly offers (the solution which resembles Kraemer's) belies its adequacy. For it is plain that the essence of Butler's problem is to explain, *if possible*, why the obvious parallelism between the two cases does not make it paradoxical for us to characterize Brown's action in the dice game case as *not intentional under the description 'throwing a six'* and yet his action in the gun case as *intentional under the description 'killing Smith'*. (My own view is that this *is* paradoxical.) To explain this it is obviously quite irrelevant to point to the relativity of intentionality to a description: one has rather to show that, contrary to our initial and untutored intuitions, the parallelism does not require us to regard as the counterpart of Brown's throwing a six in the one case his killing Smith in the other. And this is indeed precisely what Peacocke does attempt to show (in his own words, that 'the alleged problem simply arises from not comparing like with like'), using for this purpose an argument which, like Kraemer's, appeals to the notion of 'control' (though unfortunately without any explanation of this vague but crucial term). However, I shall begin by concentrating on Kraemer's account, as being somewhat the clearer of the two (I shall return to Peacocke's later).

Apparently, what makes Brown's action in the dice game not intentional under the description 'throwing a six' is (as Kraemer puts it) that Brown 'has no control over whether he throws a six. For he has no control over whether *the die that he has thrown* will turn up six' (my emphasis). By contrast, what makes Brown's action in the gun case intentional under the description 'killing Smith' is (again as Kraemer puts it) that Brown 'has some control over whether or not he kills Smith. For we assume that he has control over whether *he pulls the trigger*' (my emphasis again). But in spite of the obscurity which attaches to the notion of 'control', it is clear enough that this contrast of Kraemer's is unfair, since it compares a Brown who *has already thrown the die* with a Brown who *has not yet pulled the trigger* (and notice that Kraemer himself clearly regards throwing the die and pulling the trigger as counterparts in the two situations). To preserve complete parity between the two cases Kraemer really ought to say *either* that Brown 'has no control over whether he kills Smith. For he has no control over whether *the trigger that he has pulled* will discharge a bullet killing Smith' *or* that Brown 'has some control over whether or not he throws a six. For we assume that he has control over whether *he throws the die*'. Either way, the parallel between the two cases is perfectly preserved and no contrast emerges. In each case, whether or not Brown has some control over the final result depends upon how much of the action we take to have been completed (though to my mind the notion of 'control' is really altogether too obscure to be very useful).

Perhaps it will be felt that I have been a little too short with Kraemer, who does after all take some pains to point out that the fact that Brown 'has control over whether or not he will throw the die does not imply that he has some control over what the throw will be'. However, all that this amounts to is that the fact that Brown has control over whether or not he will throw the die does not imply that he has some control over whether *having thrown the die it will turn up six*. But if this is a reason for saying that Brown cannot have thrown a six intentionally, then there is an equally good reason for saying that Brown cannot have killed Smith intentionally. For it is equally true that the fact that Brown has control over whether or not he will pull the trigger does not imply that he has some control over whether *having pulled the trigger it will discharge a bullet killing Smith*. (If anyone feels urged to remark that Brown also has control over whether or not he aims the gun at Smith, it may be pointed out that wherever in the foregoing we speak of Brown's pulling the trigger we could equally well have spoken of his aiming the gun at Smith and pulling the trigger; this would just have resulted in a more long-winded account. But I have omitted this refinement only because Kraemer has.)

Incidentally, is it in any way significant (as the title of Ross's article suggests) that Brown *loads* the gun but does nothing comparable with the die? I think not, for in the first place it makes no difference if we suppose that Brown merely discovers the gun already loaded with a single bullet, and in the second place it makes no difference if we suppose that Brown starts with a blank die and paints six spots on one face before he throws it. Another suggestion might be that a significant feature of the gun case is that in it Brown exercises a certain *skill* (inasmuch as he has to aim the gun at Smith accurately), whereas in the dice game case he exercises no comparable skill (or at least only the minimal skill required to get the die on the table). But this may be dismissed on the grounds that nothing is sigificantly altered if we suppose Smith to be at point-blank range, so that Brown's requisite skill is minimal.

So much for Kraemer's account. It remains only to be seen that equally cogent criticisms apply to Peacocke's. According to Peacocke, the reason why Brown shoots Smith intentionally is that he 'has control over the conditional that if the gun fires, then it fires in one direction rather than another' even though he 'does not have control over whether the gun fires at all'. And the reason why Brown does not throw a six intentionally is that although he 'has control over whether. . . the die falls on one side of the table with the six uppermost *if* it falls with the six uppermost at all, . . . he does not have control over the antecedent of this conditional'. So we are asked to consider the following two conditionals and their antecedents:

(1) The die falls with six uppermost.
(2) If the die falls with the six uppermost, then it falls with the six uppermost on one side of the table rather than another.
(3) The gun fires.
(4) If the gun fires, then it fires in one direction rather than another.

Brown is said to have 'control' over (2) and (4) but not over (1) and (3). But if Brown's lack of control over (1) and possession of it over (2) imply that Brown did not intentionally *throw a six* but did intentionally *throw the die on one side of the table rather than another*, and his lack of control over (3) implies that Brown did not intentionally *fire the gun*, then what, by strict analogy, does Brown's possession of control over (4) imply? One's first thought might be that it implies that Brown did intentionally *fire the gun in one direction rather than another* (namely, at Smith). This would suit Peacocke's purposes admirably, since if Brown fired the gun at Smith intentionally, then, given as we are that he hit him, we certainly seem entitled to say that Brown shot Smith intentionally. But in fact this first suggestion is plainly erroneous. For if it were correct, and Brown's lack of control over (3) and possession of it over (4) implied that Brown did not intentionally *fire the gun* but did intentionally *fire the gun in one direction rather than another*, then given that Brown's lack of control over (1) is taken to imply that Brown did not intentionally *throw a six*, strict analogy would require us to take Brown's possession of control over (2) to imply that Brown did intentionally *throw a six on one side of the table rather than another*, rather than merely that Brown did intentionally *throw the die on one side of the table rather than another*. But this implication is clearly unacceptable: Brown did not intentionally throw a six *at all.*

What, then, *does* Brown's possession of control over (4) imply by strict analogy? Just this, I suggest: that Brown did intentionally *aim the gun in one direction rather than another* (namely, at Smith). But the fact that Brown did this intentionally no more implies that he *shot Smith* (or even fired the gun at him) intentionally than the fact that Brown intentionally threw the die on (say) the left-hand side of the table implies that he threw a six on the left-hand side of the table intentionally.

Peacocke claims that 'Butler's problem arises only when we compare a condition over which the agent does have control in the gun case with a non-corresponding condition in the dice case over which he does not have control'. We can now see that this just is not so. What corresponds to Brown's shooting Smith is his throwing a six on (say) the left-hand side of the table, and Brown has control over neither. What corresponds to Brown's throwing the die on the left-hand side of the table is his aiming the gun at Smith (and pulling the trigger, one may add), and Brown has control over both. So if we are to conclude that Brown did not intentionally throw a six on the left-hand side of the table, then strict analogy requires us also to conclude that Brown did not intentionally shoot Smith.

As I have made clear, it is my own view that the parallelism between Butler's two cases really does render our intuitive ascriptions of intentionality in these cases paradoxical. But perhaps this may be made more evident by imagining yet another case. Suppose that instead of an ordinary die Brown were to throw one specially constructed as a bomb designed to explode if and only if it came to rest with the six uppermost (but still properly balanced), and that Brown threw this die hoping that it would explode and knowing that if it did so Smith would assuredly be killed. If that happened, should we say that Brown exploded the bomb and killed Smith intentionally but not that he threw a six intentionally? (Remember that Brown explodes the bomb if and only if he throws a six and that, similarly, he kills Smith if and only if he throws a six.) I suspect that we *would* say precisely this, unreflectively at least: but I cannot see that we would be justified in doing so.

© E. J. Lowe 1980

JUSTICE, TRUTH AND NOZICK

By BERNARD MAYO

NOZICK'S 'Entitlement Theory' of justice in distribution of 'holdings' has it that *whatever arises from a just situation by just steps is itself just.* Just steps are *justice-preserving,* and are to be compared illuminatingly with correct rules of inference which are *truth-preserving.* Equally illuminating, Nozick says—and here he makes a virtue of necessity—is where the analogy fails:

> That a conclusion could have been deduced by truth-preserving means from premisses that are true suffices to show its truth. That from a just situation a situation *could* have arisen via justice-preserving means does *not* suffice to show its justice. The fact that a thief's victims voluntarily *could* have presented him with gifts does not entitle the thief to his ill-gotten gains (*Anarchy, State and Utopia,* Blackwell 1974, pp. 151–2).

He concludes that justice (in holdings) is *historical*: a major thesis subsequently developed in his disinction between *historical principles* and *end-result principles.*

I

I shall argue that the analogy between justice and truth is not one which works up to a point and then interestingly fails. It does not work at all, and Nozick's mistake is to misidentify another analogy which lies close at hand, which does not break down at the crucial point, and which can illuminate Nozick's own notion of justice-transfer much more effectively than the broken-backed truth analogy does.

Consider the first sentence of the passage quoted above. Is the term 'conclusion' within the scope of the conditionalization? If it is, then we are speaking of the truth of what might have been a conclusion, but is not actually one. But that case is not like the case of justice transfer, where it is an actual, not a counterfactual, situation we are supposed to be considering. If, on the other hand, the term 'conclusion' is not part of the conditional, but refers to an actual proposition about which we hypothesize a counterfactual way of arriving at it, then it is not a *conclusion*, and so again the analogy fails at the start. Further, a counterfactual conditional can hardly *show* anything, since such conditionals are invariably at the wrong end of such relations as showing, implying or being evidence for. "Subjunctive information" (to use the term favoured by Nozick a few lines later) does not serve to establish matters of fact. You cannot show that Q is true by means of the "fact" that it *could* have been deduced by a correct rule R from a true premiss P: you have to *deduce* Q from P, and the "facts" about what you *could* deduce trivially follow.

Thirdly, it is Nozick's claim that a combination of just acquisition and just transfer *constitutes* the justice of the ensuing situation. But a combination of true premiss and correct rule of inference certainly does not constitute the truth of a conclusion. To say that the deduction 'suffices to show' the truth of Q is a truism, but to say that the process by which a situation arises 'suffices to show' the justice of that situation ought, on Nozick's view, to be a misleading understatement.

II

I conclude that the truth-preservation analogy is a non-starter as well as a non-finisher, but before offering Nozick a better alternative, let us ask how he would have fared if he had not been tempted by the truth analogy in the first place, but had gone instead for the analogy of a board game which is explicitly, if briefly, invoked in the same passage, where 'the legitimate means of moving from one distribution to another' are redescribed as '"moves" specified by the principle of justice in acquisition'. This analogy would indeed have worked, and would have continued to work without breaking down at the crucial point. Whatever board position arises from a legitimate starting position by legitimate moves is itself a legitimate position. But that a given position *could* have been reached by legitimate (and legitimacy-preserving) moves from a legitimate original position does *not* suffice to show its legitimacy. Like the case of the thief and his victims, someone might cheat at chess and gain a commanding position by surreptitiously moving his Queen illegally; but even if the position so reached is in itself a legal one (not one in which illegality can be read straight off the board, like two white bishops on white squares), it would not, in actual play, be legitimate. Moreover the culprit would not be *entitled* to his advantageous position. Legitimacy in chess positions is, just like Nozick's justice, historical.

III

But let us return to the field of truth and inference. We have seen that the only way of establishing the truth of Q on the basis of P and R (premiss and rule) is actually to perform the deduction; to show that something could have been deduced, or can be deduced, is to deduce.

Now of course Q may be true even though someone has arrived at it in a non-accredited way. And, if it is, there must be some other, accredited way (such as R) which was not in fact used, though it could have been. In the case of justice, Nozick rejects the analogy at this point; it is not possible for a situation S2 to be just if it has arisen by an unjust process from S1. But if Nozick is right in his Entitlement Theory, he rejects the analogy for the wrong reason. He should not have selected *truth* at all as the relevant point of comparison with justice. On the other hand, if he is wrong about the Entitlement Theory, there may be room for the truth analogy after all. I proceed to develop both points.

IV

When we move from a proposition P to another proposition Q by means of a valid rule of inference R, there are at least three claims that can be made about what is preserved in the process. Only one of these claims is about *truth*. We can also ask about preservation of *belief*: must someone who believes P and moves to Q via R, believe Q? And we can ask a third question about preservation of *knowledge*: whether someone who knows P and moves to Q via R, knows Q. Let us ask each of these three questions in turn.

(1) Is R truth-preserving? The answer is certainly Yes, by definition of 'valid rule of inference'.

(2) Is R belief-preserving? The answer to this is problematic. It is well known that logical implication is not belief-preserving: none of us believes everything that is in fact implied by what he does believe. But a rule of inference is not the same thing as a logical implication; and if someone believes P, and moves by means of R to Q, it is rather hard to see how he can fail to believe Q. But fortunately we do not here need to decide this question of belief-preservation.

(3) Is R knowledge-preserving? Here the answer has to be Yes. If someone knows that P, and moves via R to Q, then he knows that Q: since Q is true (by (1)) and, if he believes Q, he is entitled to believe it, since he reached it by a correct rule of inference. Notice that even if he does not believe Q—the possibility dubiously left open by (2)—he is still *entitled* to believe it: so that knowledge will still be preserved even in the extreme case of knowledge without belief, if this *is* possible.[1]

In so far, then, as valid rules of inference are knowledge-preserving, they provide a consistent and continuing analogy with Nozick's rules of justice-preserving. And the knowledge analogy does not break down at the point where Nozick thinks the truth analogy does. For, while it is possible for Q to *be* true even though it was arrived at by a non-accredited way from P, it is not then *known* to be true. Nozick's mistake is to select truth as the property preserved, when he should have selected knowledge. Knowledge is, above all, being entitled to believe; and belief-entitlements depend on a "historical" process of transfer just as entitlements to 'holdings' do. Causal theories of knowledge, for instance, explicitly recognize that knowledge transfer depends on a historical process. And there is an extensive encroachment of the vocabulary of politics into the field of knowledge—rights, legitimacy, entitlement and authority are not functioning here as metaphors.

---

[1] See Colin Radford, 'Knowledge—by Examples', ANALYSIS 27.1, October 1966, pp. 1–11.

V

Let us, finally, ask what would have to be the case if the analogy with truth were to hold, as well as the analogy with knowledge. Perhaps justice pairs with knowledge and some other, closely associated, concept pairs with truth. Nozick conceded the breakdown of the original justice/truth analogy because Q can be true even if it was not derived from P by a valid rule, whereas S2 will not be just if it was not derived from S1 by a just process. We have just seen that the matter of how propositional items are derived is a matter of knowledge rather than of truth. But could it nevertheless be the case that S2 is just—or something similar—even though it was derived by an unjust process, as Q can be true even though it was derived by defective reasoning and so cannot then be *known* to be true?

The analogy with truth might hold if there exists, in addition to the process which *constitutes* justice in S2, a second way of evaluating S2 independently of the process by which it arose. This second way is, of course, the way envisaged by the Entitlement Theory's rival, the End-Result Theory. Those who believe in the iniquity of inequity[2] are applying end-result principles. So, even if a distribution of holdings arises by unjust steps, and is itself therefore unjust by the Entitlement Theory, yet it might still be, if not just,—shall we say?—"fair". (My scare quotes are an acknowledgment that I do not intend to discuss justice and fairness here, but only to re-examine the alethic analogies.)

More convincing, of course, and much more important to the End-Result theorists, is the opposite case: a distribution arising by just steps, and therefore itself just by the Entitlement Theory, might be "unfair". But this means that the "fairness"/truth analogy would fail after all, since "fairness" is not preserved. This is a direct result of the supposition that "fairness" was a dimension of evaluation entirely independent of justice. But truth and inference are only partly independent. Inference does not create truth, but it does preserve it. (The independence of justice and fairness is, in any case, wholly unplausible.)

---

[2] I owe this phrase to Mr Christopher Coope.

I conclude that only the knowledge analogy will give Nozick what he wants; the truth analogy will not serve the Entitlement Theory and it will not help the End-Result theorist either.

University of St. Andrews © BERNARD MAYO 1980

COUNTING PROPER CLASSES

By JAMES ROBERT BROWN

WHEN a system is proposed as a foundation for mathematics it is usually subjected to a number of tests. The criteria for the adequacy of a foundational system include some rather vague and subjective ingredients such as: the system should be "natural", "elegant", have a "small" primitive base, etc. It should also be able to capture "almost all" of classical mathematics. As well as these, there is the somewhat more objective criterion of consistency. Naïve set theory, as is well known, fails in this respect.

Long ago, Russell and Frege both insisted that a foundation for mathematics should not only tell us what numbers are, but should also explain to us just how mathematics is applied. In the case of arithmetic this means telling us just how counting is done. To simply postulate the Peano Axioms, as the Formalists did, was objectionable (not only, as Russell would put it, because it got by theft what the Logicists were trying to get by honest toil, but also) because it gave no explanation of how we could count things like apples. This, we see, amounts to the invocation of an additional criterion of adequacy for foundational systems. To be acceptable a system must give an account of, not only pure, but applied mathematics as well.

Not everyone concurs with this. Quine, for example, thinks that counting is independent of any particular account of number that we may care to give. Doubtless he is correct when he says:

> . . . Russell can be answered. . . . We can define 'there are *n* so-and-so's' without ever deciding what numbers are, apart from their fulfillment of arithmetic. That there are *n* so-and-so's can be explained simply as meaning that the so-and-so's are in one-to-one correspondence with the numbers up to *n* (*Ontological Relativity and Other Essays*, p. 44).

The criterion insisted upon by Frege and Russell does seem a bit too strong. But while I think that we should not ask the question of a foundational system: 'How does applied mathematics work?' (though it would be nice to get an answer), I do think that we must insist that applied mathematics continue to work. As far as arithmetic goes, a foundation for arithmetic must permit the counting of anything which we can intuitively count. If it does not, then regardless of its other virtues, it is not acceptable as a foundation. This shall be formulated as the counting criterion:

> (CC) *A system is acceptable as a foundation for arithmetic only if it is compatible with the counting of any objects which can be counted using unfounded, or unreconstructed arithmetic.*

For example, if it is correct to say there are five apples on the table using unfounded arithmetic then it must remain correct using arithmetic reconstructed by, say, Zermelo-Frankel set theory. Should this not be the case then the alleged foundation must be dismissed. Arithmetic must not lose any of its applicability as a result of being rationally reconstructed.

This principle, (CC), can be generalized. First, not only counting, but all of the usual applications should remain when mathematics is given a foundation. (Velocity should still be a derivative, etc.) Second, there are those who do not think that mathematics requires a foundation or they think that set theory is not such a foundation. Nevertheless, they would have to admit that there are such things as sets of apples, sets of vectors, etc. So, some theory of sets is required. (CC) could then be modified to say that a theory of sets is acceptable only if it is compatible with the counting of things which are intuitively countable. (Note that 'countable' as used here or anywhere in this paper does *not* mean 'denumerable'.) Since I am concerned with set theory as a foundation for arithmetic, (CC) will be left as it stands. The point to be made below affects set theory whether considered as a foundation or not.

Without any doubt at all, the best account of counting is set theoretical. As Quine put it above, 'That there are *n* so-and-so's can be explained simply as meaning that the so-and-so's are in one-to-one correspondence with the numbers up to *n*.' This means that there is a function which maps the set of so-and-so's onto a set of numbers. The important thing to note is that the so-and-so's are in a set.

So far I have postulated two seemingly innocuous principles: one, that rationally reconstructed arithmetic must be as powerful in application as unreconstructed arithmetic; and two, that the correct account of counting is set theoretical. This would be quite uninteresting were it not for the fact that there are formal systems parading around as the foundations of mathematics which are not up to scratch. Those systems which run afoul of these two principles include the set theories of Kelly-Morse (KM), von Neumann-Gödel-Bernays (NGB), and Quine's Mathematical Logic (ML). The common feature of these systems is, of course, the distinction that they make between set and proper class.

DEFINITION: A is a *set* iff $\exists B: A \in B$
A is a *proper class* iff A is not a set.
(Cf. Monk, *Introduction to Set Theory*, p. 14, which is a presentation of KM.)

We should note that set theories are presented with or without individuals (*Urelemente*). Here we assume that provisions have been made. In this definition, A is not an individual.

From this definition we have the following

THEOREM: A, is a proper class iff $\sim \exists B: A \in B$.

The next thing to be considered is the fact that proper classes are intuitively countable; that is, they can be counted by unreconstructed arithmetic. For instance, we can let V be the class of all sets, *Ord* be the class of all ordinals, and *Card* be the class of all cardinals. These are all proper classes. It is true to say that three proper classes have just been defined. (I must here stress that I am making no claims about the cardinality of proper classes. I am only claiming that I can intuitively count the proper classes that were just defined, and that there are three of them.)

Another way of putting the matter is to claim that the following question makes sense. 'How many specific proper classes does Suppes refer to on p. 12 of *Axiomatic Set Theory*?' The answer, which can be found by turning to the appropriate page, is two. Thus, from a pre-analytic point of view proper classes can be counted. Unfounded or unreconstructed arithmetic can be applied to these objects.

If we make the following four assumptions disaster will strike.

(1) (CC) holds.
(2) The correct account of counting is as above; i.e., there is a one-to-one function from the set of things counted onto a set of numbers.
(3) KM proper classes are intuitively countable.
(4) KM is the correct set theory. (This involves two things: one is that, in virtue of (2), the sets used to count are KM sets; the other is that proper classes actually exist—whatever is meant by mathematical existence.)

Of course it is quite easy to see why there is a problem. If KM is the correct set theory then the one-to-one correspondence involved in counting is a KM function which maps one KM set onto another KM set. If KM proper classes are counted they have to be members of KM sets. According to the above theorem this cannot be the case. However, KM proper classes are intuitively countable; thus, (CC) is violated. The same difficulty arises with NGB, ML, and any other set theory which countenances proper classes.

The difficulty completely vanishes if we drop (4) and say that ZF is the correct set theory. Now no absurdity arises from counting KM proper classes simply because they are members of ZF sets. There is no inconsistency in this. We must not think that because proper classes are countable they must exist. Countability might be a necessary condition for existence; it certainly is not a sufficient one. We can count unicorns; I dreamed of three of them last night.

According to some theories of meaning anything at all that we may say about unicorns will be false. Accordingly, on these same theories, anything said about proper classes will be false because they do not exist. Fine; I will only claim that we can count proper classes in the same sense that we can count unicorns.

If we insist that a set theory, to be a candidate for the foundations of mathematics, comply with (CC), then out go the likes of KM, NGB, and ML. Of course their interest as formal systems remains. Nothing that I have said touches on their internal consistency, only on their truth-value. (I am speaking as a realist in calling them 'false'. Others may prefer a different term.)

For quite some time now there has been little or no debate about which set theory is correct, nor about which of ZF, KM, etc. is *the* foundation of mathematics. (Of course queries about the Axiom of Choice, the Continuum Hypothesis, etc. are of a different nature.) Indeed, such questions are usually dismissed for the very good reason that the various set theories all have the same mathematical consequences, more or less. It works out that $5+7=12$ in all of them. But while they may all have the same *pure* mathematical consequences, they do not all have the same *applied* mathematical consequences. By utilizing (CC) in the appraisal of competing set theories we can make some rational choices. (CC) does not yield a unique set theory, but it does separate the sheep (ZF and others without proper classes) from the goats (KM and others with proper classes). The field of acceptable candidates is considerably narrowed since any set theory which postulates the existence of proper classes gets into trouble. And so it follows that proper classes cannot exist and that any set theory which posits them cannot serve as a foundation for arithmetic.

For their helpful comments I wish to thank W. Harper, K. Okruhlik, and G. Pearce.

University of Western Ontario © James Robert Brown 1980

WHAT IS, NECESSARILY IS, WHEN IT IS

By C. J. F. Williams

IN a recent article[1] I claimed that in the famous ninth chapter of *De Interpretatione* Aristotle puts forward for our consideration an argument of the following form: if tomorrow it will be the case that *p*, it was true yesterday that in two days' time it would be the case that *p*. But nothing that can now be done can alter what was already the case yesterday. So if it will be the case tomorrow that *p*, we cannot now prevent its being the case tomorrow that *p*. I further claimed that nothing Aristotle says in the rest of the chapter constitutes a rebuttal of this argument. Mr. Malcolm Lowe (Analysis 40.1, January 1980) questions this latter claim, and draws our attention to Aristotle's remarks at 19a 23–27. These, he believes, are intended as the rebuttal of the above argument, which occurs at 18b 9–16, whereas the sentences which follow (19a 27–b 2) are intended to rebut the argument from the Law of the Excluded Middle (18a 34—b 9).

It was my contention that the argument of 18b 9–16 crucially employs the notion of truth; and since this notion plays no part in the argument of 19a 23—27, it is difficult, though, as we shall see, not impossible, to see how the latter can be regarded as a rebuttal of the former. But what *is* going on at 19a 23–27?

In passing at 19a 27 to his consideration of the argument from the Law of the Excluded Middle, Aristotle says 'And the same account holds . . .'. What account? Lowe has an answer to this which I find less than convincing. The correct answer, I believe, is as follows.

Aristotle half-diagnosed the fallacy which we now call 'operator shift'. It is what his followers talked about under the heading 'fallacy of composition'. And it is a version of this fallacy which Aristotle is exposing in 19a 27–b 2.[2] The fatalist, he is claiming, confuses two possible interpretations of 'One or other of them is necessarily the case' ('*thateron anankaion*': 19a 29). A more regimented language would distinguish this from 'Necessarily, one or other of them is the case'. The ambiguity is exactly parallel to that of 'If something is the case, it is necessarily the case', which needs to be distinguished from 'Necessarily, if something is the case, it is the case'. This latter source of ambiguity Aristotle shows himself aware of elsewhere (cf. *De Caelo* I xii, 281b 4, sqq.). Now it is apparent that another case of the same ambiguity is to be found in the first sentence of 19a 23–27, 'What is, necessarily is, when it is'. Does this mean 'What is, necessarily, when it is, is' or 'What is, when it is, necessarily is'? As often in Plato and Aristotle, the word literally translated 'is' here would more naturally be rendered 'is the case'. How then are we to provide a formal representation of 'What is the case, when it is the case, is necessarily the case' and its fellow? To do this most conveniently let us adopt one of Prior's suggestions for a symbolism for metric tense logic using Fnp for 'It will be the case n time-intervals from now that p', and allowing n to range over negative as well as positive numbers. Fnp, for $n<0$, will then mean the same as 'It was the case m time-intervals ago that p' (otherwise, Pmp) where m is the corresponding positive number. $F0p$ $(=P0p=p)$ will mean 'It is now the case that p'. The only metric tense operator we shall need is Fn[3].

---

[1] 'True Tomorrow, Never True Today', *The Philosophical Quarterly*, 28 (1978), pp. 285–99.

[2] Cf. the use of the word '*dielonta*' at 19a 29 with '*dielōn*' at 166a 26 (*De Sophisticis Elenchis*, ch. IV), from which the scholastic '*sensu diviso*', no doubt, derives.

Armed with this symbolism, we may represent 'What is the case, necessarily, when it is the case, is the case' by $\Pi n\Pi pLCFnpFnp$ and 'What is the case, when it is the case, necessarily is the case' by $\Pi n\Pi pCFnpFnLp$. Since $F0p$ is an instance of Fnp, $\Pi pCF0pF0Lp$ is an instantiation of the latter of these universal propositions; and since $F0p=p$, it is equivalent to $\Pi pCpLp$, that is, 'What is the case, necessarily is the case'. But this is precisely what Aristotle says is to be distinguished from 'What is, necessarily is, when it is'. If, therefore, we move from $\Pi n\Pi pLCFnpFnp$ to $\Pi n\Pi pCFnpFnLp$ by an operator shift, we move from there to $\Pi pCpLp$ by impeccable logic. If, at whatever time it is the case that p, it is then necessarily the case that p, it does indeed follow that if it *is* the case that p, it is (now) necessarily the case that p. 'The same account' can indeed be given of the move from 'What is the case, necessarily, when it is the case, is the case' to 'What is the case, necessarily is the case' as can be given of the move from 'Necessarily one or other is the case' to 'One or other is necessarily the case'. The essential move in each case is fallacious and the fallacy is that of shifting the operator. Aristotle would be entitled to say that the same account could be given of each.

The move from $\Pi n\Pi pLCFnpFnp$ to $\Pi pCpLp$ is fallacious, but it is not clear at this stage that it leads to error. 'What is, necessarily is, when it is', even when interpreted as $\Pi n\Pi pCFnpFnLp$, is not obviously false. It seems to be a statement of the principle that what is now the case or has been the case is necessary, in the sense of being now unpreventable, generalized to include the corollary that what will be the case, when it in turn *is* the case, will be *then* unpreventable. It does not have the undesirable consequence that what will be the case is even now, before it is the case, unpreventable. Or does it? As we have seen, $\Pi pCpLp$ follows from $\Pi n\Pi pCFnpFnLp$ as an instantiation of it, in virtue of the fact that $F0p=p$. $\Pi pCpLp$, if it is interpreted as saying that whatever *is* the case is now unpreventable, looks harmless. But $\Pi pCpLp$ seems to entail $\Pi pCFmpLFmp$, for $m>0$, by legitimate substitution. And this is precisely the consequence which we found undesirable, namely that what will be the case is even now, before it is the case, unpreventable. In virtue of the equivalence of $F0p$ and p, $\Pi pCFmpLFmp$ is equivalent to $\Pi pCF0FmpF0LFmp$, and it is the possibility of inserting L between the two metric tense operators in the consequent which causes the trouble. Against this someone may object that the inference of $\Pi pCFmpLFmp$ from $\Pi n\Pi pCFnpFnLp$ is an instance of another operator-shift fallacy. But that would have been so only if we had derived $\Pi pCFmpLFmp$ directly from $\Pi pCFmpFmLp$, a substitution instance of $\Pi n\Pi pCFnpFnLp$. In fact we derived $\Pi pCFmpLFmp$ from another substitution instance of $\Pi n\Pi pCFnpFnLp$, namely $\Pi pCF0pF0Lp$, by deletion of $F0$ and substitution of Fmp for p. Since these moves are legitimate we seem to be required to reject $\Pi n\Pi pCFnpFnLp$, since it leads to the undesirable $\Pi pCFmpLFmp$. The fallacious inference of $\Pi n\Pi pCFnpFnLp$ from $\Pi n\Pi pLCFnpFnp$ does after all lead to error.

---

[3] Cf. A. N. Prior, *Past, Present and Future*, Oxford: At the Clarendon Press, 1967, p. 97.

$\Pi n\Pi pCFnpFnLp$ seemed acceptable as the expression of a generalization of the thoroughly Aristotelian thesis that what is now the case or has been the case is necessary, in the sense of being now unpreventable. If it is in fact unacceptable, perhaps this is because it is an inadequate symbolization of the principle which we find intuitively sound, rather than an indication that the principle itself is unacceptable. What is required is, rather, that limitations be placed on legitimate substitutions for p in $\Pi n\Pi pCFnpFnLp$. In virtue of our allowing n in Fnp to range over all numbers, negative as well as positive, any substitution for p in $\Pi n\Pi pCFnpFnLp$ will be equivalent to some proposition of the form Fmp. The rule needed for substituting for p will be that such substitutions are always to be made in the form of an explicitly tensed proposition, i.e., in the form Fmp, and that they shall only be legitimate if m is $\leqslant 0$. This will allow us to say, for instance, that if it will be the case tomorrow that there is a sea battle then, it will be the case tomorrow that it is necessary then that there is a sea-battle. It will not allow us to say that if it is the case now that there will be a sea-battle tomorrow, it is now necessary that there will be a sea-battle tomorrow; nor will it allow that, if it was the case yesterday that there would be a sea-battle in two days' time, it was necessary yesterday that there would be a sea-battle in two days' time. Interpreted in accordance with this rule, what $\Pi n\Pi pCFnpFnLp$ says is, after all, true.

We cannot seriously suppose that Aristotle was aware of all this. Nevertheless, it is difficult not to feel that 'the same account' which, he says, he is giving both in 19a 23–27 and 19a 27–b 2 is the claim which we should express by saying that the operator-shift fallacy had been committed in each case. Can his remarks in 19a 23–27 be understood in this sense?

Change one word in 19a 23–27, and the answer is, 'Obviously, yes'. Change 'when' to 'if' at all its occurrences in this passage, and you get

Aristotle making a distinction between 'Necessarily, if *p*, then *p* ($\Pi pLCpp$)' and 'If *p*, then necessarily *p* ($\Pi pCpLp$)'.[4] To confuse these two is, as has been remarked, to commit the operator-shift fallacy. Can we suppose, then, that Aristotle is failing to notice the difference between 'What is, when it is, necessarily is' and 'What is, if it is, necessarily is'? This ought not to be ruled out as an impossible account of what is going on in Aristotle's mind.[5]

But how, in this case, could 19a 23–27 be considered relevant to 18b 9–16? The fallacious move from $\Pi pLCpp$ to $\Pi pCpLp$ is not only illegitimate but also productive of error. If it is confused with a move from 'What is, necessarily, when it is, is' to 'What is, when it is, necessarily is', this latter move might also be construed as a passage from truth to falsehood. 'What is, when it is, necessarily is' can be interpreted, as we have seen, as an expression of the principle, suitably understood, that what is the case, when it is the case, is then unpreventable. This principle seems to imply that if it was the case (was true) yesterday that in two days' time there would be a sea-battle, it was necessary yesterday, in the sense of being then unpreventable, that in two days' time there would be a sea-battle: which is the basis of the argument of 18b 9–16. If Aristotle in 19a 23–27 conceives of himself as denying this principle, confused with 'Everything that is necessarily is', he may think of himself as there refuting the basis of the argument of 18b 9–16. On the face of it he seems to be affirming 'What is, when it is, necessarily is' and denying 'Everything that is necessarily is ($\Pi pCpLp$)'. But perhaps he thinks that it is only in the sense of 'What is, necessarily, when it is, is' that 'What is, necessarily is, when it is' can be truly affirmed, and that it is only by confusing this with 'What is, when it is, necessarily is' that anyone would come to believe the latter. This would be reinforced if he supposed that this was the same mistake, not merely the same sort of mistake, as that of confusing 'What is, necessarily is, if it is ($\Pi pLCpp$)' with 'Everything that is necessarily is ($\Pi pCpLp$)'.

It is difficult to be sure that by interpreting Aristotle in this way one has established a connection, of the kind Lowe alleges, between 18b 9–16 and 19a 23–27. It is a great deal easier to feel that Aristotle is aware of a scope ambiguity in 'What is, necessarily is, when it is', as he is in 'One or other of them is necessarily the case', and that this is what makes him say he has given 'the same account' of each. But this not very plausible theory of the connection between 18b 9–16 and 19a 23–27 is not the only consideration which could lead me to a recantation of my previous view that the argument of 18b 9–16 is nowhere in this chapter rebutted by Aristotle. What I had in mind when I put forward that view was something on these lines: On one interpretation of what Aristotle is doing in the chapter, he is maintaining that while, necessarily, every proposition is either true or false, only propositions about the past or present are either necessarily true or necessarily false. He regards the fatalist view as the result of the operator-shift fallacy, by which we are tempted to pass from the truth that, necessarily, every proposition about the future is either true or false, to the falsehood that every proposition about the future is either necessarily true or necessarily false. But this provides no refutation of the argument of 18b 9–16, which relies on the premiss that some proposition, not about the future, but ostensibly about the past, e.g., that what was said yesterday was true, is either necessarily true or necessarily false. On the other main interpretation, Aristotle's position is that the Law of the Excluded Middle does not apply to propositions about the future. But this again seems to leave the argument of 18b 9–16 untouched. For to all appearance the proposition 'What was said yesterday was true' is a proposition about the past, and so must itself be either true or false—Aristotle, at least, gives us no reason for thinking differently. It seems that on neither main interpretation of what Aristotle is doing in Chapter IX has he produced any refutation of what is argued in 18b 9–16.

---

[4] Aristotle's actual words, when amended in this way, are equivalent to 'What is the case necessarily is the case if it is the case' and 'What is the case, if it is the case, necessarily is the case', which is merely a pleonastic version of 'What is the case necessarily is the case'.

[5] Dorothea Frede has claimed that in the *Analytics* Aristotle often uses the Greek word for 'when' in a sense indistinguishable from that of the Greek word for 'if'. She cites 25b 32, 26a 9, 26a 39, 26b 34. Cf. '*Omne quod est quando est necesse est esse*', in *Archiv für Geschichte der Philosophie*, vol. 54, 1972, pp. 153–167, n. 3.

A different view, however, is held by Professor Ackrill,[6] in his commentary on this chapter, as I should have remembered. He regards Aristotle as challenging, not only, or not so much, the Law of the Excluded Middle, as applied to contingent future propositions, but the Principle of Bivalence. That is to say, we cannot attribute a truth value to such a proposition. And if we cannot now say that it is true that there will (or will not) be a sea-battle tomorrow, we could not yesterday have said that it was true that there would (or would not) be a sea-battle in two days' time. A proposition about the future, if contingent, does not have a truth value at a time when the state of affairs whose existence it states is still future. And so 'What was said yesterday was true' and 'What was said yesterday was false' may well neither of them be true (perhaps both are false). The argument from past truth to necessity of the future does not get started.

It now seems to me that Ackrill's account of how Aristotle thinks himself to have refuted the argument of 18b 9–16 is better than either Lowe's or my own tentative one, sketched above. But it leaves unanswered the question, 'What *is* Aristotle up to in 19a 23–27?'

University of Bristol

© C. J. F. WILLIAMS 1980

---

[6] J. L. Ackrill, *Aristotle's Categories and De Interpretatione* (Clarendon Aristotle Series) Oxford: At the University Press, 1963, pp. 140–2.

RATIONALITY AND INTRANSITIVE PREFERENCES

By R. I. G. HUGHES

IT is clearly impossible to construct a linear preference ordering of three alternatives, A, B and C, from an agent's pairwise choices among them, if those choices yield an intransitivity; that is, if, while preferring A to B and B to C, he prefers C to A. Perhaps for this reason, many writers on the topic consider such an intransitive set of preferences as conclusive evidence of irrationality.[1] A fairly simple example will show that this conclusion is mistaken.

Consider an informed voter, wondering which candidate for some municipal office he should support. Three local figures, Aley, Baley and Caley, have entered the lists. With regard to their civic policies he has no difficulty in ranking them. Politically Baley stands further away from him than Aley, and Caley further away still, so that if their announced programme were all that concerned him he would rank them in alphabetical order. However, he is also concerned with their corruptibility; he has good reasons for believing Aley to be more venal than Baley, and Baley more so than Caley, so that, were honesty the only consideration, the reverse order would result.

All that has been shown so far is that the voter has a problem. He could perhaps solve it by constructing two interval scales: on one of them each candidate is assigned a score which depends on his announced programme, and on the other a score depending on his degree of venality. Then, provided that the voter can come up with some way of equating the intervals on the two scales, all he needs to do is to add the two scores obtained by each candidate. The result would be a new composite interval scale on which all three candidates are ranked; clearly, just by virtue of being a cardinal scale, this would display transitivity. The slight artificiality of the procedure does not make it incoherent. However, the voter may object to it on the following grounds. He may say that the construction of an interval scale based on corruptibility would not reflect his attitude to this issue: for him its importance only appears when a significant distance separates two candidates. When, in terms of honesty, two candidates are less than some critical distance apart, he neglects the difference between them; on the other hand, should they be widely separated the question of integrity becomes paramount, overriding all other considerations. It may well be that the distances between Aley and Baley, and between Baley and Caley, are both less than this critical distance, while that between Aley and Caley exceeds it. Thus, given the choice between Aley and Baley, he would vote for Aley on the basis of his political programme, since, although Baley is the more honest, the difference between them is small enough for that consideration to be unimportant. Similarly he would vote for Baley over Caley. However, faced with the choice between Aley and Caley, he would find the difference in corruptibility too great to ignore; this difference would override his purely political preferences and he would vote for Caley.

---

[1] For a sample from recent years, see the definition of a 'rational player' in Anatol Rapoport, *Two Person Game Theory, The Essential Ideas*, Ann Arbor (University of Michigan Press), 1973 (p. 34), or the discussion of transitivity in Moshe F. Robinstein, *Patterns of Problem Solving*, Englewood Cliffs, New Jersey (Prentice Hall), 1977 (p. 323), or that in Richard G. Niemi and William H. Riker, 'The Choice of Voting Systems', *Scientific American*, June 1976, v. 234, no. 6. Luce and Raiffa (in *Games and Decisions, Introduction and Critical Survey*, New York (John Wiley), 1957) are more careful. They offer rather vague hints suggesting that when intransitivities arise, the incommensurability of factors affecting choices may be at work. The example I give is not of this kind: I assume that the voter is able to compare the factors affecting his choice.

We see that the ordering relation expressed by his pairwise choices is not transitive, and so does not offer a useful way of ranking all three candidates. Can we convict him of irrationality? He has told us the factors which would influence each of the three choices, and the algorithm which would determine them; his method of selection is perfectly systematic. We might say that the way in which corruptibility enters the calculation is itself evidence of irrationality. Notice that this would be nearly, but not quite, to beg the question at issue: we would be demanding that a rational agent exhibit a transitivity of *indifference* with respect to each factor that influenced his choice. The fact that many people fail to do so, and instead exhibit "catastrophic" changes of preference when one of these factors exceeds a certain threshold, by no means guarantees the rationality of such behaviour; whether in general it should be thought of as irrational is a large issue which I will leave to one side.[2] To convince us that it *need* not, and in the present case that it *should* not be so regarded, the voter may argue as follows.

Assume there are two candidates for an elected office. Should the more corrupt of the two be elected, the damage done by his own dishonest actions would probably be less than that due to the resulting loss of morale in the public service. Thus what is important is not so much the venality of that candidate as the probability of such a reaction to his election. This probability in turn increases abruptly if it is clear to those in the public service that the electorate is being given a real choice, in other words, if there is a markedly more scrupulous candidate available. The choices made by the voter reflect his estimates of the likelihood of this reaction: the question of integrity becomes an issue for him precisely when this likelihood increases abruptly, that is, when the difference between candidates becomes significant.

---

[2] That was the (somewhat hesitant) conclusion of Davidson, McKinsey and Suppes, in 'Outlines of a Formal Theory of Value, I' (*Philosophy of Science*, vol. 22, no. 2, April, 1955), p. 145.

Now the voter may be wrong in his account of the dynamics of this reaction, but that does not make him irrational. Even if he is correct we may think that this shows irrationality to be rife in the public service; but again, to take account of the irrationality of others is not to betray it in oneself, rather the reverse. He has perfectly good reasons for making the choices he does.

Agents may or may not argue in the way I have described: the aim of this note is not to give an empirical account of how intransitivities occur, but to argue that they need not arise through a failure of reason on the part of the agent. The conclusion obviously has implications for the programme of constructing complete utility scales for information about the pairwise ordering of alternatives. What the voter in this example would do if the electoral procedure required him to rank all three candidates in order (or just to make a first choice from among them) is not clear; what we cannot assume (even if we take him to act with perfect rationality) is that this ordering could be determined from the choices he would make were he to consider the candidates two at a time.

Princeton University © R. I. G. HUGHES 1980

REFERENTIAL SHIFTS

By Rod Bertolet

GARETH EVANS has suggested that the history of 'Madagascar' makes trouble for the causal theory of names sketched by Kripke. That I won't dispute, but I do want to dispute what seems to be a presumption some have about how we should attempt to account for change of reference or denotation such as the one that occurred with 'Madagascar'. (Evans' criticism is in 'The Causal Theory of Names', in S. Schwartz (ed.), *Naming, Necessity, and Natural Kinds* (Cornell, 1977), pp. 192–215; for Kripke's sketch see 'Naming and Necessity' in D. Davidson and G. Harman (eds.), *Semantics of Natural Language* (Reidel, 1972), pp. 253–355 & 763–769.)

To review the example briefly, Marco Polo picked up the name 'Madagascar' from natives who used it to pick out part of the African mainland, but he mistakenly thought it was a name for the island now known as Madagascar, and so used it. Since our own uses derive from Marco Polo, it would appear that the view Kripke offers will carry us along a causal chain leading back to a portion of the mainland, making *it* the denotation of our uses of the name. Since this doesn't happen to be true, causal theorists need to augment their account in some way.

With this much I agree; what I do not agree with is the way in which some philosophers think the explanation of this shift should, and perhaps must, go. I would claim to find this view latent in various writers, but Michael McKinsey has been explicit about it—one piece of evidence McKinsey offers in support of his own theory of names (the details of which need not concern us here) is that it makes it 'easy to explain how it was possible for Marco Polo's very first uses of "Madagascar" to denote the island' ('Names and Intentionality', *The Philosophical Review* 87 : 2, p. 199). He seems, then, to find it at least desirable that such an explanation be forthcoming, and it is this that I wish to question.

First, though, we might ask why this might be thought virtuous. Presumably, to make room for intuitions we have about truth. Suppose that upon his return to Italy Marco Polo had remarked 'Madagascar is bigger than Sicily' (and for convenience suppose that the natives' Madagascar was smaller than Sicily); we are I think inclined to say that his remark was *true*, and that it is not merely virtuous but *necessary* to explain this by saying that 'Madagascar' denoted the island. For *how else* could we explain the truth of Marco Polo's remark? (Notice that this argument, which I have heard in discussion but not seen in quite this form in print, promotes McKinsey's suggestion from a desideratum to an adequacy condition on any explanation of the shift.) What I am suspicious of is a presumption that I think lies behind the How else? question;

it is this presumption, which I will uncharitably call the semantical preoccupation, that I shall try to describe and attack.

Borrowing a piece of strategy from Evans, let's start with common nouns instead of names. In the 14th Century those things in the extension of 'dungeon' were the strongholds of castles; by the 19th Century, the extension of 'dungeon' consisted of dark subterranean prisons. It is doubtful if this shift came about by royal decree—the most plausible explanation for it is that people started calling unlit underground prisons dungeons, with this usage gradually taking hold. Somebody (just call him Charles) had to start the change, possibly by remarking something like 'Egbert was put in the dungeon for poaching the king's deer', and we would probably say that this remark was true. And we might then expect an argument to the effect that we must therefore say that Charles' use of 'the dungeon' denoted some particular prison.

What—but all that—makes this persuasive is the presumption that what is true is *the sentence* 'Egbert was put in the dungeon . . .': the semantical preoccupation. Intuitively, this seems wrong to me. Though Charles did make a remark that was true, or in some sense say something true, the sentence he uttered, I think, was not true. This intuition can be backed up with an argument. Charles, after all, made a mistake: the prison into which Egbert was thrown was not, at the time of his utterance, in the extension of 'dungeon'. Charles' use of 'the dungeon' *could* not denote a prison, subterranean or otherwise. (I am relying on what I take to be the standard assumption that for some term t to stand in the semantical relation of denoting x, x must be in the extension of t.) The only way I see around this is to claim that the utterance *puts* the prison in the extension of 'dungeon', but if we were to say this we would be unable to account for the fact that Charles did make a mistake. So the conditions of the problem—the fact that it is a *change* of denotation (more broadly, extension) that we are to account for, and one that took place over five centuries—seem to preclude us from saying that the sentence Charles uttered was true. And if this is so, the semantical preoccupation is a mistake. This result would appear to leave us in a quandary concerning the intuitions about truth noted earlier; I will return to it shortly.

The situation concerning 'Madagascar' is, I think, parallel. Marco Polo was *wrong* about the name of the large island east of Africa, and it would behove us to preserve this fact in our account of the shift in the denotation of 'Madagascar'. We could, of course, just add that the island got its current name through a mistake, but this is more than a little *ad hoc*. (It is interesting, if not conclusive, that if Marco Polo had been advised of his error by the natives, he would very likely have stopped calling the island 'Madagascar', for the very good reason that 'Madagascar' was not its name.) While 'Madagascar' clearly became the name of the island, it is not obviously virtuous, let alone necessary, to incorporate the claim that Marco Polo's first uses of the name denoted the island into our account of this change.

But now what about the intuitions concerning truth mentioned earlier? We want to say that Marco Polo said something true about the island, and we do not want to say that he said something false about part of the mainland, when he uttered 'Madagascar is bigger than Sicily'. And doesn't my suggestion preclude us from endorsing this intuition? No. Not unless we endow the intuition with some theory that it does not obviously demand. It is, I think, a consequence of some well known work of Donnellan's that we can say something true *of* something even if we use a term that bears no semantical relation to that thing (K. S. Donnellan, 'Reference and Definite Descriptions', *The Philosophical Review* 75 : 3, pp. 281–304; it is not clear that Donnellan agrees with this conclusion). While Donnellan was originally concerned with definite descriptions, I think the same point holds for common nouns and names. It is perfectly possible to say that Marco Polo said something true of the island by uttering 'Madagascar is bigger than Sicily' *and* to deny that the sentence is true.

More precisely, a more careful description of what I am calling the semantical preoccupation would be this: the inclination to promote our intuition that something true is (in some sense) said into the claim that the sentence used to say it is true. My criticism of this inclination is that all our intuitions really tell us is that something true is said *of* something, and that moving from this to a claim about sentences constitutes an infusion of a good deal of theory that our intuitions will not bear out. Once we stop being horrified by the notion that a speaker might (again, in some sense) say something true by using a sentence not itself true, the motivation for making the jump to sentences evaporates.

Those whose suspicions about the claim made in the previous paragraph run deep might find it useful to compare a case in which an error is made but where no shift in reference ensues. Suppose that Marco Polo's sole misuse of 'Madagascar' occurred in an attempt to engage the natives in conversation—pointing to the island, he uttered 'Madagascar sure is a big island'. Further suppose that the natives corrected the error, so that no other mistakes occurred. There seems to be a strong intuition to the effect that *Marco Polo said something true of the island,* but no temptation at all to say that *the sentence Marco Polo uttered is true*. While this example is contrived, I think that many (if not all) actual cases in which speakers use the wrong name or noun for something deserve a similar description.[1]

Finally, I would like to tie my somewhat vague suggestions about how to explain the shift in denotation of 'Madagascar' to a general account of linguistic change. The rough idea is that this semantical shift, and others like it, occurs gradually, and is mediated by the (I think) nonsemantical relation of speaker reference. Speaker reference is part and parcel of speaker meaning, and speaker meaning seems to be a plausible candidate for the explanation of change of meaning for various parts of speech. (The examples I will discuss are perhaps controversial, and you should feel free to substitute.) Sticklers who consult dictionaries would insist that 'aggravate' means make some condition more serious, with the result that 'Carl's attitude really aggravates me' is ungrammatical. It isn't. But if we presume the accuracy of lexicographers, that sentence used to be ungrammatical, and 'aggravate' formerly did not mean irritate. Now it does. Why?

---

[1] I wish to thank L.-M. Russow for prompting this paragraph.

The *only* plausible explanation I can see is that speakers mistakenly used 'aggravate' to mean irritate, that they were understood to mean irritate, and that this usage gradually caught on so that 'aggravate' *now means* irritate (since some still respect the earlier usage, perhaps it is ambiguous). A similar story must, I think, be told about the transformation of 'hopefully' into a sentential modifier, at least in American English. If I am right in thinking this is a generally applicable account of at least some kinds of linguistic change, sketchy though it may be, we should expect it to have a place in the change concerning 'Madagascar', since that change also proceeded via an error. How a causal theorist might make room for this sort of account with a greater degree of precision is of course another question.

(Bibliographical note: While the suggestion of the previous two paragraphs is at least close to an even vaguer one made by Kripke at the end of 'Speaker's Reference and Semantic Reference' (in P. French *et. al.* (eds.), *Contemporary Perspectives in the Philosophy of Language* (Minnesota, 1979), pp. 6–27), my own fondness for it stems from conversations with Lloyd Eggan in 1975–76; I also owe the dungeon example to Eggan.)

Purdue University

© Rod Bertolet 1980

SCREENING-OFF AND THE TEMPORAL ASYMMETRY OF EXPLANATION

By A. David Kline

GIVEN three causally related and temporally ordered events, where *a* occurs before *b* and *b* before *c*, our intuition is that generally *b* ought to be explained in terms of *a*, not *c*. This intuition could be rationalized, if one adopted the blatant temporal constraint that events be explained by reference to events that temporally precede them. Most philosophers frown upon this move for reasons that need not concern us here. It is thought that a more promising alternative would be to show that some basic non-temporal restriction on explanation guarantees that at least most events be explained by temporally prior events.

One of the exciting claims Wesley Salmon makes for his statistical relevance analysis of explanation is that it accounts for the temporal asymmetry of causal explanation. In particular Salmon believes that the screening-off rule can function as the needed basic restriction.

I shall be examining Salmon's argument in the remainder of this paper. Since his argument is very dense at its critical junctures, the first task will be to render the structure of the argument more perspicuous. When this is done the argument, at least *prima facie*, contains a contradiction. The contradiction can be removed, and in a way sympathetic to Salmon; but no longer does screening-off account for the temporal asymmetry of explanation in Salmon's straightforward way. Yet there is a sense in which screening-off is relevant to the problem. In locating this sense the central role of manipulation in Salmon's argument will become evident—a role he has not sufficiently emphasized.

I

I assume that the reader is familiar with the statistical relevance analysis but I remind him of some of the concepts central to the present discussion. The screening-off rule is as follows:

> When one property in terms of which a statistically relevant partition in a reference class can be effected screens off another property in terms of which another statistically relevant partition of that same reference class can be effected, then the screened-off property must give way to the property that screens it off (1, p. 55).

A property C constructs a statistically relevant partition of A, the reference class, with respect to property B if and only if $P(A \cdot C, B) \neq P(A, B)$. (Salmon uses Reichenbach's nonstandard notation. '$P(A \cdot C, B) \neq P(A, B)$' reads 'the probability of B given A and C is not equal to the probability of B given A'.) A property D screens off property C from B in reference class A if and only if $P(A \cdot C \cdot D, B) = P(A \cdot D, B) \neq P(A \cdot C, B)$.

II

With this machinery in hand what Salmon would like to show can be seen quite straightforwardly. Given our three causally related and temporally ordered events, *a*, *b* and *c*, the claim is that in most cases *a* will screen off *c*. Hence *b* ought to be explained in terms of *a* not *c*. It is no accident or prejudice on our part that we believe that events ought to be explained in terms of temporally prior events. Our belief is consistent with a basic restriction on explanation—the screening-off rule.

Salmon 'demonstrates' how screening-off accounts for the temporal asymmetry by examining a particular example. He then suggests that in a similar way screening-off would function in most cases.

Suppose a spoiled meal is served to the members of a theatrical company who fall ill and later cancel their intended performance. Our intuition is that the illness should be explained in terms of the earlier meal of spoiled food rather than the subsequent cancellation. Salmon believes that screening-off can capture this intuition. In particular, he wants to show that meals of spoiled food, F, screen off cancellations, C, from illnesses, $L \cdot M$, (the leading lady and leading man falling ill for brevity) in the general reference class of days, A. More formally he wants to show that

$$(1)\ P(A \cdot F \cdot C, L \cdot M) = P(A \cdot F, L \cdot M) \neq P(A \cdot C, L \cdot M).$$

I shall quote, in order, parts of Salmon's argument for (1) and discuss them.

> It is not true, without further qualification, that the eating of spoiled food F screens off the cancellation of the performance C from the joint illness $L \cdot M$. Although the eating of spoiled food does make it probable that the two actors will be ill, the fact that the performance has been cancelled supplies further evidence of the illness and makes the joint illness more probable (1, p. 73).

I have two comments on these sentences. Screening-off was defined as a relation between properties. Salmon has slipped into treating screening-off as a relation between event types. But I see no serious problem with this move at least in the present context.

More importantly, one ought to be surprised by these opening sentences. If we are to take them seriously, it would appear to be tantamount to abandoning Salmon's strategy. He is saying that C is statistically relevant, i.e.,

(2) $P(A{\cdot}F{\cdot}C, L{\cdot}M) \neq P(A{\cdot}F, L{\cdot}M)$.

If (2) is true, (1) cannot be true. Hence it appears that screening-off does not apply.

Since this is an important matter and one to which we shall return it is worth emphasizing what Salmon has admitted. Let us define a mere observer as one who observes instances of a certain causal sequence without interrupting, interfering or participating in that sequence. This notion could be rendered sharper but for the present purposes it will do. What Salmon is admitting is that a mere observer with respect to the causal sequence F, $L{\cdot}M$, C will be able to discover that proposition (2) is true.

But Salmon does not see this as the end of the matter. He continues his argument.

> At this point experiment must be admitted. It is clearly possible to arrange things so that the play goes on, illness or no illness, by providing substitutes who never eat with the regular cast. Likewise, it is a simple matter to see to it that the performance is cancelled, whether or not anyone is ill (1, p. 73).

These two sentences are the crucial part of Salmon's argument. From them he concludes

(3) $P(A{\cdot}F{\cdot}C, L{\cdot}M) = P(A{\cdot}F{\cdot}\bar{C}, L{\cdot}M)$,

and in turn

(4) $P(A{\cdot}F{\cdot}C, L{\cdot}M) = P(A{\cdot}F, L{\cdot}M)$.

It is obvious why he wants to establish (4) since his goal is to prove (1). But what is the argument for (4)? (Disregard the fact that (4) contradicts (2); we shall take that up later.)

What he is suggesting is that no matter whether $L{\cdot}M$ is the case or $\sim(L{\cdot}M)$ is the case it is within the power of the experimenter to bring about C or $\sim C$. Interpreting this idea more modestly and more to the point we can imagine an experimenter *manipulating* C and $\sim C$ and computing the probability of $L{\cdot}M$. The experimenter does not simply observe the frequency of C and $L{\cdot}M$, and $\sim C$ and $L{\cdot}M$ as a mere observer would do. He actively manipulates C in an effort to affect the frequency of $L{\cdot}M$. Putting it formally, he manipulates C and gathers the following statistics:

(5) $P(A{\cdot}F{\cdot}C, L{\cdot}M)$
(6) $P(A{\cdot}F{\cdot}\bar{C}, L{\cdot}M)$

Salmon is quite convinced that the performance of such an experiment will result in (3) which again entails the needed (4). Salmon is correct, but notice that it is a bit odd to describe the above as an 'experiment'. After all we can tell *a priori* what the results will be, viz., (4). Our *a priori* confidence comes from the fact that at the time one manipulates C, $L{\cdot}M$ has already occurred or failed to occur. Furthermore we believe that the experimenter is free in his choice to bring about or prevent C. So how could C be statistically relevant to $L{\cdot}M$?

Though this line of reasoning lies behind our confidence, and I believe Salmon's confidence, in (3), it is a real virtue of his view that he need not rely on it. If anyone is not convinced, he can do the experiment and compute (5) and (6) and see for himself. Salmon need not assume anything about free will or one's knowledge of $L{\cdot}M$ prior to manipulating C.

What follows is the last element of Salmon's argument.

> If we approach the common meal of spoiled food in the same experimental fashion, we can easily establish that it is not screened off by the cancellation of the performance. We can arrange for the two leading actors to have no meals supplied from the same source and ascertain the probability $P(A{\cdot}\bar{F}{\cdot}C,\ L{\cdot}M)$. This can be compared with the probability $P(A{\cdot}F{\cdot}C,\ L{\cdot}M)$ which arises under the usual conditions of the two actors eating together in the same restaurants. Since they are not equal, the common spoiled food is statistically relevant to the joint illness, even in the presence of the cancellation of the performance, that is,
> (7) $P(A{\cdot}F{\cdot}C, L{\cdot}M) \neq P(A{\cdot}C, L{\cdot}M)$ (1, p. 73)

(I numbered the above statement for future reference.)

The new experiment is the analogue of the previous one. This time F is manipulated, brought about and prevented, and the probability of $L{\cdot}M$ is computed. More precisely the following probabilities are determined:

(8) $P(A{\cdot}F{\cdot}C, L{\cdot}M)$,
(9) $P(A{\cdot}\bar{F}{\cdot}C, L{\cdot}M)$.

As the result of this experiment, Salmon is confident that one will see that (10) is the case and hence (7) is the case.

(10) $P(A{\cdot}F{\cdot}C, L{\cdot}M) \neq P(A{\cdot}\bar{F}{\cdot}C, L{\cdot}M)$
(7) $P(A{\cdot}F{\cdot}C, L{\cdot}M) \neq P(A{\cdot}C, L{\cdot}M)$

Again knowledge of (7) is dependent on the experiment but it is quite obvious that Salmon's belief is correct. But now given (7) and the previously demonstrated (4) we can derive (1)—what Salmon wants to show. F does appear to screen off C from $L{\cdot}M$.

The difficulty which must now be faced is the contradiction alluded to earlier. Since (2) and (4) are contradictory, one of them must be given up. The argument for (4) was persuasive and necessary if (1) is to be defended. But (2) seems equally evident.

The problem resides in an ambiguity. C and F were originally used to represent event types—cancellations of performances and eating of spoiled food. When experiment was resorted to, C and F represented sub-classes of the original event types. At that point they represented cancellations that were brought about by the experimenter and the eating of spoiled food that was brought about by the experimenter. Let us call these sub-classes c and f. If we pay attention to this ambiguity both (2) and (4) can be true. C is statistically relevant to $L{\cdot}M$ but c is not.

Since F does not screen off C from $L{\cdot}M$, one cannot claim that it accounts for the temporal asymmetry of explanation. It is not the case that one ought to explain $L{\cdot}M$ in terms of F rather than C because F screens off C. The latter is false.

Salmon must be understood as claiming that the fact that f screens off c accounts for the temporal asymmetry in explanation. But now there is a gap in the argument. If F were statistically relevant to $L{\cdot}M$ and C were not, it is easy to see why $L{\cdot}M$ ought to be explained in terms of F. But granting that f is statistically relevant and c is not, why should $L{\cdot}M$ be explained in terms of F?

The gap can be filled by adopting some version of the manipulability theory of causation/explanation. $L{\cdot}M$ ought to be explained in terms of F since the explanandum event ought to be explained in terms of events that can *in principle be manipulated* so as to affect the probability of the explanandum. Events of the F type are such events. Events of the C type are not such events. We come to see this by the experiments, by seeing that f screens off c.

Though I am sympathetic with the manipulability theory as an account of the temporal asymmetry problem, it has not been my intention to defend it but rather to show its central role in Salmon's argument. It is important to realize that in the role of mere observer no relevant asymmetry emerges. As long as we simply collect statistics in the presence and absence of $L{\cdot}M$ no asymmetry emerges. One must actually manipulate instances of F and C in hopes of bringing about $L{\cdot}M$ if he is to discover an asymmetry.

Iowa State University © A. DAVID KLINE 1980

REFERENCE

(1) Salmon, W., 'Statistical Explanations', in *Statistical Explanation and Statistical Relevance*, edited by W. Salmon, University of Pittsburgh Press, 1971.

MATERIALISM AND THE UNITY OF CONSCIOUSNESS

By ANDREW WARD

BY consulting memory, one can be aware of the contents of many past experiences united in one act of consciousness. But what—in this sense of the phrase 'the unity of consciousness'—makes the unity of consciousness possible? One answer which is now in favour, although claims along these lines have periodically occurred for centuries, is that there is good reason for thinking that the unification is made possible by a continuously existing brain (thus providing the connection between the past experiences and one's present consciousness of them). What reason is there for thinking that the material world exists? One answer which is presently fashionable, although similar replies may well go back at least to Descartes, is that given the actual order of perceptions amid the flux of experiences, the best hypothesis is that the material world does exist. (An excellent example of a contemporary writer who is prepared to give both these answers is Mr. J. L. Mackie: see his *Problems from Locke* pp. 62–71 and pp. 199–203).

I shall contend that one cannot justifiably claim that the brain makes the unity of consciousness possible unless one's grounds for believing in the existence of matter do *not* rely on a best hypothesis argument. And I should add that I have never seen a materialist give the vestige of a respectable case for believing in the external world that does not depend, either wholly or partly, on this attempted proof.

Now it is a presupposition of the attempted proof of a material world that memory-experience is trustworthy. For the argument rests on the claim that, if there were *not* a material world, it would be highly unreasonable to have expected the succession of perceptions to have come in the order it actually has (since, without a material world, it would be an inexplicable coincidence, if out of all the possible combinations of successive perceptions, successive perceptions should actually be so *similar* as to suggest that they are observations of persisting objects); whereas, if there *were* a material world, it would be wholly reasonable to have expected the successive perceptions to be similar (since the continuous existence of something external, to cause the succeeding perceptions, explains the similarity between them). From which it is concluded that, given the *actual* order of perceptions, there is good reason to believe in a material world. And obviously the trustworthiness of one's present memory-experience has to be presupposed in order for one to be justified in claiming that past perceptions *have* come in any given order. But by employing precisely the same type of argument, it follows that, if past experiences of given times are *not* connected with one's present ostensible memory as of those past times, it would be highly unreasonable to expect one's present memory-experience to be trustworthy (since, without a connection, it would be an inexplicable coincidence if, out of all the possible ostensible memories as of those past times, the present memory-experience should be of just such a content as actually to *match* the experiences of those past times); whereas, if there *were* a connection, it would be wholly reasonable to expect the present memory-experience to match the past events, i.e., to be trustworthy (since the continuous existence of something, to store and reproduce the content of the past experiences, explains the matching between them and the present memory-experience). From which it may be concluded that, given one's memory-experience *is* trustworthy, there is good reason to believe that past experiences are connected with one's present consciousness of them. (I am not saying that the attempted proof of matter is valid. In fact I think it is invalid: see my 'Direct and Indirect Realism', *American Philosophical Quarterly* 1976, and 'What's not really wrong with Phenomenalism', *Australasian Journal of Philosophy* 1973. The point is that anyone who does accept the argument with respect to matter ought *also* to accept it with respect to memory).

Of course, if it *were* highly unreasonable to trust the accuracy of memory-experience, it would be correspondingly highly unreasonable for advocates of the best hypothesis argument to claim to have good reason for believing in a material world. Consequently their attempted proof must presuppose that past experiences are connected with one's present consciousness as of the past. But there can be no good reason for thinking that what makes the unification possible is material because, for the attempted proof to have any plausibility, the accuracy of memory-experience must evidently be assumed *logically prior* to discovering anything about the *actual* order of the series of perceptions; and, therefore, *logically prior* to having proved (as it is claimed) the existence of the material world. Nor will it do *later*—i.e., *after* one has thereby purportedly proved the existence of matter—to claim to have good reason for thinking that the brain makes the unification possible. Since, intentionally or not, one would then be claiming that, if the brain did not exist, it would be highly unreasonable to trust memory-experience. But this means that one's attempt to prove the existence of matter will need to presuppose the existence of matter (the brain), and so will be circular. In other words, in claiming to have good reason for thinking that the brain makes the unification possible, one would, in effect, be claiming to have no good reason for thinking that matter, and hence the brain, exists.

I have not *denied* that the brain makes the unity of consciousness possible. My position (like the materialists that I oppose) has been epistemological; that one cannot hope to *prove* that the unification is made possible by the brain unless one's grounds for believing in matter do not themselves presuppose that the very thing thought to be secured by the brain must already have been secured.

University of York © Andrew Ward 1980

COUNTERFEIT COINS AND FORGED PAINTINGS: CAVEAT EMPTOR

By MICHAEL WREEN

IN a well-known passage, Gilbert Ryle writes:

> A country which had no coinage would offer no scope to counterfeiters. There would be nothing for them to manufacture or pass counterfeits of. They could, if they wished, manufacture and give away decorated discs of brass or lead, which the public might be pleased to get. But these would not be false coins. There can be false coins only where there are coins made of the proper materials by the proper authorities.
>
> In a country where there is a coinage, false coins can be manufactured and passed; and an ordinary citizen, unable to tell which were false and which were genuine coins, might become suspicious of the genuineness of any particular coin that he received. But however general his suspicions might be, there remains one proposition which he cannot entertain, the proposition, namely, that it is possible that all coins are counterfeits. For there must be an answer to the question 'Counterfeits of what?' (*Dilemmas*, Cambridge, 1953, pp. 94–95).

Virtually the same argument has been advanced by John Passmore (*Philosophical Reasoning*, New York, Basic Books, 1961, p. 108), and, more recently, by Jay Rosenberg. Rosenberg's main concern, however, is forged paintings:

> . . . 'it could be the case that paintings are always forgeries.' . . . [this] is false. For a forged painting is a *copy* of some original painting, and it could not be the case that all paintings were copies. If *all* paintings were copies, no paintings would be originals, but if no paintings were originals, there would be nothing for the supposed copies to be copies *of* (*The Practice of Philosophy*, Englewood Cliffs, N. J., Prentice-Hall, 1978, p. 19).

Even Colin Grant, who is chary of polar concepts arguments, is willing to invest in counterfeit coins:

> [C]ounterfeiting is defined *simpliciter* as copying coins,

he says, and this is why

> There are no . . . uses of 'counterfeit' which could give the expression 'universal counterfeiting' a meaning ('Polar Concepts and Metaphysical Arguments', in *Clarity Is Not Enough*, H. D. Lewis, ed., London, Allen and Unwin, 1963, p. 272),

with this apparently meaning that to say that all coins are counterfeits is to utter a meaningless sentence or to assert a necessary falsehood. Other remarks of Grant's also suggest an acceptance of the counterfeit coins argument *and* a willingness to buy Rosenberg's forged paintings:

> The skeptic's statement 'the senses always deceive us' is more like 'All paintings attributed to Vermeer are false' than it is like 'All coins are counterfeits'. In my proposed analogy I do not mean by 'false,' 'fake,' i.e., a forgery intended to deceive experts and the public about their authorship, for in that case Ryle's question 'Fakes of what?' would arise (p. 272).

Despite the pervasiveness and apparent force of this argument, I believe that it is unsound. In the case of coins, the claim in question is that, given that there are coins, not all coins are counterfeits. Here is a counter-example to that claim, and I will follow up my counter-example with a counter-argument.

A country just coming into existence decides it needs a coinage. Designs are drawn up, plates made, dies cast, a mint built, and newspaper stories, complete with artists' sketches of the new coins, released to the public. But the day before the coins are scheduled to go into circulation and just hours before the first coins are to be pressed, the plates are stolen—and, perhaps, destroyed. The next day coins passably similar to what the government had planned to manufacture and circulate are manufactured and circulated by a gang of counterfeiters (not necessarily the thieves). All coins would then be counterfeits. (See D. W. Hamlyn's *The Theory of Knowledge*, New York, Doubleday, 1970, pp. 18–19, for a similar counter-example.)

A similar case could be constructed as a counter-example to the claim that it's impossible that all paintings are forgeries. For instance, at a given time there could be no paintings. Someone could then announce that tomorrow he will unveil his (yet unpainted) paintings, which, he says, look like such-and-such. The budding painter is then kidnapped, Van Meegeren goes to work, and the next day paintings having all the properties specified in the hopeful painter's announcement are unveiled and exhibited, the aspiring painter's name having been signed on each painting and labels attributing the paintings to the thwarted painter appearing by each painting's side. All paintings would then be forgeries.

Counterfeit coins are coins which are not genuine but which are represented as genuine, with the intention to deceive; and similarly with forged paintings. If my counter-examples are acceptable, there need not be any actually existent coins or paintings which counterfeits or forgeries are copies of. However, my counter-examples also suggest that, in Hamlyn's words, 'The pattern is important' (p. 19), that is, there must be some specifications regarding design or organization and materials, and/or some plate or mould or template, in order for there to be counterfeits or forgeries. But this is not so. I will first offer counter-examples to this claim and then go on to explain the notions of a counterfeit and a forgery at greater length.

A newly forming government could announce that on a certain day coins will be distributed; but out of laziness, sheer stupidity, tangled red-tape, or corruption, it could have no coins minted on that day, no dies cast, no plans drawn up, and not the least idea of what the coins are to look like or what their size, shape, chemical composition, etc., are to be. Counterfeiters could go to work in such circumstances; they could produce and distribute coins on the announced day, probably taking in the public in so doing.

With paintings, there is really no need to appeal to an imaginative counter-example. Just as a matter of empirical fact, very few forged paintings are either copies of existing paintings or painted from specifications, verbal or otherwise: in most cases they are simply paintings executed in the style or manner of a particular artist, school, or period. Take, for example, Van Meegeren, one of the most famous forgers. Van Meegeren didn't copy a single painting of Vermeer's nor did he work from verbal descriptions of putatively existent or possibly existent paintings of Vermeer. What he did was paint in the style or manner of Vermeer. ('Style' and 'manner', it should be noted, *may* have to be construed very broadly in some cases of forgery so as to include such things as subject matter—though it should also be remembered that, in the case of Van Meegeren, it was the very fact that the subject matter of his forgeries *differed* from known Vermeers that enabled him to perpetrate a successful deception.) True, there is a Vermeer style, and this provides somewhat of a pattern, or something pattern-like, to work from or to copy. But first of all, the notion of a pattern has been extended a good bit here; and secondly and much more importantly, an extremely ingenious forger could 'invent' a new artist, school, or period, complete with characteristic style or manner, right on the spot. This might be extremely hard to do convincingly in some cases—the 'creation' of Antarctican art of the tenth century B.C. would be quite difficult—but it is never a logical impossibility. And indeed it is not even always 'extremely hard to do convincingly': James Macpherson's 'Ossian' poems, purporting to be by a third-century Gaelic poet, were forgeries, in a broad sense of the term; but although no such poet existed and there is virtually no Gaelic literature of that period, Macpherson succeeded in deceiving large numbers of people. The pattern is not important—at least as far as the *concept* of a forgery is concerned.

I have said that counterfeit coins are coins which are not genuine but which are represented as genuine, with the intention to deceive; and similarly with forged paintings. If what has been said above is correct, the norm of genuineness does not have to do with copying or answering to a particular pattern, however loosely 'pattern' is construed. What then is the norm, and why have Ryle, Passmore, Rosenberg, Grant, and, to a lesser extent, Hamlyn, been misled? (In all fairness to Ryle it should be noted that Ryle himself does not use the term 'copy', although, except for Hamlyn, all of the other philosophers mentioned above do.)

In answering the first of these questions I would like to begin by focusing on forged paintings and borrowing rather heavily from Nelson Goodman. Goodman holds that a work of art is 'autographic if and only if the distinction between original and forgery of it is significant; or better, if and only if even the most exact duplication of it does not thereby count as genuine'. Thus, 'painting is autographic, music non-autographic, or *allographic*' (*Languages of Art*, Indianapolis, 1968, p. 113). The important point here, however, is that 'Authenticity in an autographic art always depends on the object's having the requisite, sometimes rather complicated, history of production' (p. 119n). This means that 'essential to the genuineness of a painting is the history of production; the only way of ascertaining that the *Lucretia* before us is genuine is thus to establish the historical fact that it is the actual object made by Rembrandt' (p. 116). If Goodman's account is correct, an account of a forgery follows rather naturally: a forgery is a painting which does not have, but is represented to have 'The requisite, sometimes rather complicated, history of production', and is so represented with the intention to deceive. And if so, it is perfectly possible for all paintings to be forgeries.

It is at least arguable that a slightly different account of counterfeit coins would have to be given. For a strictly parallel account would have it that a counterfeit coin is a coin which does not have, but is represented to have the requisite, sometimes rather complicated, history of production, and is so represented with the intention to deceive. But, it could be said, this does not seem to be so. Imagine a powerful nation discovering a small and technologically primitive island civilization. Suppose that at the time of discovery each and every native islander wears a necklace on which a number of small but qualitatively indistinguishable circular discs are strung. If the more powerful nation were to conquer the island civilization, one thing it might do is declare a new coinage: genuine coins are all and only those small circular discs which were worn by the natives the day the more powerful nation took over.

Here it would seem that the genuineness of the coins has nothing to do with the production or history of production of the small circular discs: neither the conquering nation nor the natives may know anything about or care at all about the history of production of the discs. In fact, discs indistinguishable from and having the same history of production as genuine coins, e.g., discs carved five hundred years ago from a boulder on a nearby island, would be regarded as counterfeits if they were introduced into the native economy. Monetization here depends, and always depends, the argument would conclude, on the requisite, sometimes rather complicated, institutional (i.e., governmental or quasi-governmental) connections. Concerns with the history of production are dispensable in the case of coins; the concept of a genuine coin is almost wholly institutional. But concerns with the history of production cannot be dispensed with in the case of paintings; although a credited painter need not himself apply paint to canvas (as credited painters did not in certain city-schools of painting during the Renaissance), and so the requisite history of production may be loosely specified and, to some extent, institutionally laden—still, concerns with the actual history of production are always relevant.

There is much to be said for this argument against a parallel account of genuine coins, and, if 'history of production' is interpreted rather strictly here, there may be nothing that can be said in answer to it. However, if a slightly broader or more generous interpretation is allowed, a powerful rebuttal is available. For if it is permissible to speak of the history of provenance, that is, the genesis of the thing, as the key notion here, then the necklace case and similar putative counterexamples present no problem for a parallel account. Whether it is small circular discs carved from a boulder or other already produced (or not produced at all, e.g., naturally occurring) objects which are appropriated by the government, it is the institutional origin and the history of the things in question which determine their genuineness. By this I mean that the question we ask of a coin represented or claimed to be genuine is: Does the coin have the proper institutional origin and history; that is, does it have a history, as a coin of the type it purports to be, with a beginning which is government sanctioned or government 'blessed'? Roughly speaking, a genuine coin is one whose government-issue birth certificate and whose history, as a coin of that type, are all in order. This is very close to Goodman's account of an authentic painting.

But why then have Ryle, Passmore, Rosenberg, Grant, and Hamlyn gone wrong? I think that the answer to this has to do with the until now sleeping rider 'with the intention to deceive'.

Counterfeiters and forgers are not fools. They know that, as things go in the world, at least most of the time the persons they intend to deceive will not be deceived unless their wares bear a fairly strong resemblance to certain fairly precisely describable paintings or coins already believed to exist or already believed quite possibly to exist. Given their intention and given this knowledge, counterfeiters and forgers are, in most cases, forced to produce resemblances, that is, coins and paintings based on a pattern or something pattern-like already believed—usually truly believed—to have instances (or at least an instance). The fact that they do produce resemblances is a contingent one, though, for it depends on the contingent facts that almost no one would be taken in by anything less than a strong resemblance and that counterfeiters quite rationally believe as much. Patterned paintings and coins, then, are the result of the intention in conjunction with the rational belief. And it is this upshot of the intention/rational belief pair which has been understandably but mistakenly thought to be essential to the notions of a counterfeit and a forgery.[1]

Temple University © MICHAEL WREEN 1980

---

[1] I would like to thank Elizabeth Beardsley, Monroe Beardsley, Paul Edwards, Carole Knapp, Meg Wreen, and an ANALYSIS referee for a number of useful suggestions on an earlier draft of this paper.

THE SHIP OF THESEUS

By THEODORE SCALTSAS

THE predominant opinion in current literature concerning artifact identity is that in cases of conflict, the continuity of form cannot be overruled by any other sufficiency condition for artifact identity, such as the identity of parts, while the latter can be overruled by the former. Cases such as the puzzle of Theseus' ship are presented to be non-problematic and prima facie, but not essentially, puzzling. I wish to argue that there are no grounds for accepting such an overruling hierarchy of sufficiency conditions for the identity of artifacts—where there is conflict between such conditions—and that the example of Theseus' ship as well as other similarly constructed cases are actual paradoxes.

'I want to insist that the continuity of form condition is a sufficient and non-defeasible condition' '. . . the continuity of form condition is non-defeasible, while the identity of original parts condition is defeasible' (Brian Smart, 'The Ship of Theseus, the Parthenon, and Disassembled Objects', ANALYSIS, October 1973, 34.1, pp. 24–7; hereafter, Smart-2) '. . . the existing roles of identity had proved perfectly adequate for this unusual case [a version of the Theseus ship puzzle] and had yielded a non-arbitrary and clear-cut decision . . .' (Brian Smart, 'How to Reidentify the Ship of Theseus', ANALYSIS, 32.5, April 1972, p. 148; hereafter, Smart-1). Also, 'The false beliefs . . . have been responsible for generating a puzzle [in the case of Theseus' ship] where no real puzzle or need for a decision exists.' (Smart-2, p. 27.) In 'Identity, Origin and Spatio-temporal Continuity' (*Philosophy*, forthcoming) I have argued that there are at least three major criteria that may serve as sufficiency conditions (or at least as conditions) for determining whether the identity of an object or an artifact is recovered in various instances of repair, reassembly, reconstruction, etc., of an object or artifact that has ceased to be. I claim there that the sufficiency condition might be the spatio-temporal continuity of the form of the object, or the identity of the parts, or the identity of the matter of the object, and further, that the origin and the history of the object might play an integral part in determining whether the initial object is recovered or not. What I wish to emphasize further and argue for here is that there is no sharply defined hierarchy of sufficiency conditions, so that in cases of conflict we are not always in a position to determine whether the new object is identical to the initial one or not. The reason is that the cases of conflict are so rare in everyday life that we have not had the need to compare the various sufficiency conditions and determine the relative strength of each in cases of conflict. Hence, our intuitions are blunt when it comes to making such decisions as in the case of Theseus' ship and similar puzzles, which shows these puzzles to be irresoluble on the basis of our existing intuitions concerning the identity of recovered artifacts.

I shall begin by presenting a version of the puzzle of Theseus' ship which will show that Smart's claim that the continuity of form condition is a sufficient and non-defeasible condition for reidentifying an artifact is not correct. Consider that there are two identically constructed ships that make the sacred voyage from Piraeus to Delos every year. In a particular year the head priest of the mission—who happens to have a philosophical bent—decides to exchange the planks of the two ships while they are returning to Piraeus from Delos. (For details about the actual process of exchange see Neurath's Ship.) Let us suppose that the planks of one of the two ships were painted red and the other blue. Then we can distinguish the original red ship (O–R), the original blue ship (O–B), the final red (F–R) and the final blue one (F–B). Throughout the whole trip, the form of O–R has remained spatio-temporally continuous and so has the form of O–B. The head priest, being sympathetic to Smart's convictions on artifact identity, faces no difficulty in reidentifying O–R in F–B and reidentifying O–B in F–R. The continuity of the form of each ship settles the question. Consequently, without any hesitation, he places the holy relics which were, according to the religious tradition, always kept in the special cabin of O–R in the special cabin of F–B, thus, for him, keeping them in the proper place. But when they reach Piraeus, the crowd waiting to follow the procession which would take the relics from the ship's cabin to the temple is outraged to realize the head priest's sacrilege: he had transported the relics on the wrong ship, namely on F–B. Calmly the head priest explains to them what had happened, reassuring them that no sacrilege had been committed and that the relics had always been kept on O–R, only that O–R now has new parts. But the crowd would not have any of this. They all saw very clearly that the relics came out of O–B; they do not care how many times the head priest might have exchanged the planks of the two ships back and forth; the fact remains that the ship with the red planks is O–R, which is supposed to carry the relics, and the ship with the blue planks is O–B, which is not supposed to have them. There is no way for the head priest to get out of his predicament by trying to convince them that the ship with the blue planks is actually O–R. The falsehood of his claim is evident for everyone to see. The head priest protested, but in vain; it was plain as day that he was guilty of sacrilege.

It is clear that for the crowd, the identity of the original parts of the two ships was the primary sufficiency condition, overruling the continuity of the form of each of the two ships. It must be noted that the colour of the ships or the planks had no influence in deciding which ship was which. Had the head priest, as a last resort, assured the crowd that he had not exchanged the planks of the two ships but only painted the blue ship red and the red blue, the crowd would have settled down since the relics would, after all, have been carried in the proper ship; (the colour of the ship had no religious significance, so the crowd would not mind the head priest changing it.) (In the above example we could have avoided all mention of the colour of the planks and spoken only of the original planks of each of the two ships; colours were introduced only for making reference to them easier.)

What is significant in the above example is that there is nothing the head priest can tell the crowd to convince them that the continuity of form is the primary sufficiency condition for identity, overruling the identity of parts. In Smart-2, 'This condition [the continuity of form] cannot be defeated by the identity of parts condition and it cannot be defeated by any other condition. Secondly, I want to amend the formulation of the second condition and call it "the identity of original parts condition". This still applies only when the original object has been disassembled and so cannot defeat or be defeated by the continuity of form condition' (p. 25). Why? Why must the continuity condition defeat the identity of parts? If Smart's intuitions tell him that it has to be so, the crowd's intuitions tell them that it certainly does not. The blue ship (O–B, F–B) was the ship made up from the blue planks, and that's that. The continuity of form is too abstract a criterion to fight such substantial evidence as the whole of the plank-constitution of the ship. (Needless to say that it would be hopeless to tell the crowd that F–B *constitutes* O–B but *is* O–R. Such distinctions work well—even for the crowd's intuitions—in cases such as 'this bronze statue is the bronze vase you saw in my studio last month', but not in the case of the two ships where the final product, F–B, has *identical shape* with the original one, O–B, though not spatiotemporally continuous, and the parts of F–B have *identical shape* with the parts of O–B. In this sense, it might be correct to say that in the above example the identity of parts was not the sufficiency condition but the primary condition, since the identity of the form of O–B and F–B also influenced the crowd's decision in identifying them.) It is precisely because there are no established arguments for determining which sufficiency condition overrules the other, and because our intuitions have not developed a standard hierarchy of such conditions—not having had the need to do so in everyday life—that we are confronted with the paradoxical situation in the above story, where the head priest's intuitions concerning the identity of the ships are diametrically opposed to those of the crowd, and neither of the two is more correct than the other. There is simply the possibility of being differently inclined on the subject of condition hierarchy from your fellow-men, and there is no evidence they can bring to you or you to them to convince one another who is right.

But the comparison of defeasibility of the above two conditions is not the only grey area in making decisions concerning the identity of artifacts. Consider the conditions of the identity of parts alone. By itself it can be the source of fundamental disagreement, which cannot be resolved by argument or on the basis of generally accepted positions on the subject. In Smart-1 we find the following claim: 'When a plank had been removed from X and replaced by a new plank, it was the new plank that was now part of X. The old plank *had been* a part of X but was no longer' (p. 148). One can think of many situations in which this would not be an acceptable claim to make. For example, I take a part of my bicycle out and replace it with another; the old part that I keep in my garage is a part of my bicycle. It is not on my bicycle now but it certainly is my bicycle's part. Even if I lent this part to a friend of mine to use on his bicycle, I could point at it and say: 'this is a part of my bicycle' (and mean 'is' not 'was'). There is also a sense of its being a part of his bicycle, too, but this only shows that there are different criteria for being a part and different avenues that lead to claims that x is part of y. There is no end to the complication of the 'being a part' cases one can devise to show the possibilities of conflict in deciding whether x is part of y or part of z or both, but part of z in a stronger sense than part of y, etc. Instead of enumerating such cases, I have listed below eight conditions, each of which can affect our decision on whether x is part of y or of z:

(1) x is an attached and functioning part of y;
(2) x is an attached but non-functioning part of y;
(3) x has been removed from y but is not used as a part of any other object or for any other function;
(4) x has been removed from y but it has not been replaced by any other such part;
(5) x is a part that could be fitted to y but it has not been fitted yet (e.g., if I buy a wheel for my bicycle and have not attached it to my bicycle yet, I could say—while explaining to a friend what each article in my garage is—'that's a part of my bike'; the more impending and definite the attachment of the part to the artifact is, the more we are inclined to speak of 'part of y' rather than 'part for y');
(6) x is one of the original parts of y (when y was constructed);
(7) x is originally a part of y (in x's first use as a part);
(8) x has at some point in its life span been functionally attached as a part to y; (it is a part from/of y, y', y'', . . .).

I am convinced that this is an incomplete list of the various senses in which x can be taken to be a part of y. Even now it would take a computer to figure out the possible conflicts of 'being a part' sufficiency conditions that can occur in various cases in each of which two or more of the above criteria would oppose one another; it would become intolerably complex when various degrees of power of overruling were introduced to the effect that if, e.g., x is part of y under one sense but part of z under another, the latter overrules the first and hence x counts as part of z. The main problem, of course, would be that no such uniform overruling hierarchy could be established—on the basis of our intuitions on the subject, since we have not needed to sharpen our intuitions to such a degree of accuracy in decision-making on this subject—which means that in the long run, conflicts might arise not only from the different senses of 'being a part of y' that might apply to a given situation, but also from the different hierarchies of overruling that might be employed in making the decisions. Paradoxical situations would abound! But this is not alarming; were such decisions significant in everyday life, we would have resolved the paradoxes by settling down to some specific senses of 'being a part of y' and some specific hierarchy of overruling between them. Given that we have not, though, we cannot introduce, in resolving Theseus' ship paradox and the like, preferred senses (such as Smart's 'original parts', Smart-2, p. 25) and preferred hierarchies of senses—if more than one sense of 'being a part' is acknowledged—that just happen to appeal to our own intuitions.

However, further problems might arise regarding the role that the form of the object might play in identifying the recovered artifact. One factor is the spatio-temporal continuity of the form. But the identity of forms—of the original and the final—where there is no continuity of form, is a second factor that can even come into conflict with the first (as it did in the case of the blue and the red ships when the identity of the forms of O–R and F–R combated the continuity of forms of O–R and F–B; in the case of the crowd, the identity of form between O–R and F–R was one of the conditions that affected their decision regarding the identity between O–R and F–R rather than O–R and F–B). A third form-criterion is the similarity of form; the reidentification of a restored ancient statue will depend on the similarity (though not quite identity, unless the criteria for form identity are relaxed) of the form of the restored statue to the form of the original one. (This is not meant to commit one to the 'best candidate' identity relations.) Thus, the element of form can affect our decisions in reidentifying artifacts either through the spatio-temporal continuity of form, or the identity of forms (without continuity), or through the similarity of forms.

Apart from form and parts, another element that can influence our reidentification decisions is the identity of matter. I have tried to show in 'Identity, Origin and Spatio-temporal Continuity' that identity of matter between x and y *can* result in the reidentification of x and y—when there is no form-continuity or even continuous existence of the parts of x and

y. In some cases of artifact identity, matter identity can be the primary criterion affecting our reidentifying α in a recovered artifact, though identity of the form—without continuity of the form—of the artifacts and of the form of the parts of the artifacts might also be a factor in making this decision. So, matter identity may result in x constituting y, but it may also result in x being identical with y. Further, as David Wiggins has pointed out in *Sameness and Substance* (Basil Blackwell 1980, pp. 92–94), there are the problems of deciding what exactly is meant by the covering concepts that are used in making a reidentification. We can see the problems that arise from such differences in concepts from the discussion of the Parthenon example (see Francis Dauer, 'How not to Reidentify the Parthenon', ANALYSIS, 32.2, December 1972, pp. 63–64, and Smart-2). The Parthenon, as a national treasure, can be taken to be the temple or the collection of marble slabs (Smart's distinction) or, we could further add, the relics at that particular site or simply the building (in a restored, building-like state, rather than a pile of ruins). There will not always be agreement on whether the Parthenon *is* one or the other of the above, and consequently, different identity criteria will be applied according to the decision on the precise covering concept.

What I have tried to show here is that we can create many paradoxes of the Theseus' ship type, precisely because we do not employ a sharply defined code of sufficiency conditions and overruling hierarchies between these conditions when deciding questions of artifact reidentification after some ceasing-to-be of an artifact and recovery. Discovering the lack of such a reidentification code is discovering something very significant about our intuitions and habits concerning decision making on artifact identity. But this does not mean that in the face of the lack of such a code we can assume or impose a code which we happen to find appealing and convincing; only the need for making such decisions in everyday life will force us to develop a functionally acceptable code of artifact reidentification.

Wolfson College, Oxford

© THEODORE SCALTSAS 1980

FRANK JACKSON AND THE SPATIAL DISTRIBUTION OF SENSE-DATA

By A. OLDING

IN directly perceiving, or sensing, sense-data, nothing can come between ourselves and the sense-data we directly perceive. If the colour and shape of a sense-datum are sensed then they are grasped just as they are, in themselves, with no distortion. The same is true of the relations between sense-data if these also are sensed.

It is, therefore, surprising to find that Frank Jackson, in defending the idea of sense-data as laid out in three dimensions, says, 'I am committed to this view that sense-data are at various distances from perceivers and at various angles to their line of sight' (Frank Jackson, *Perception, a Representative Theory*, Cambridge 1977, p. 102). Because Jackson believes that sense-data are in the same space as other things (and not in some mysterious 'mental space') there is an unobjectionable sense we might give to this statement. This is that sense-data are at various distances from perceivers, considered as ordinary physical objects, and at various angles to their lines of sight if these are taken as the directions in which their eyes are pointing. In this they differ not at all from a host of ordinary physical objects, perceived and unperceived.

This sense, in which *your* sense-data are just as much at various distances from, and angles to, me as they are for you is not the one that Jackson needs. Jackson quotes H. H. Price as saying, 'It is a plain phenomenological fact that visual fields have the property of depth' (H. H. Price, *Perception*, Methuen 1954), and this is true if what is meant by it is only that it is a phenomenological fact that the sense-data of a given visual field are disposed in three-dimensional space. If he wishes to go further and claim it is also a plain phenomenological fact that the visual field extends away in depth *from one* so that its contents are *sensed from* (or, perhaps we should say, sensed as sensed from) that which senses, the sensor, then he must be mistaken. This just is because in sensing there can be no privileged 'point of view'—the things sensed and the spatial etc. relations between them are sensed exactly as they are.

This is true even if, as a matter of fact, sense-data do stand in certain spatial relations to the sensor. Perhaps the sensor is a non-spatially extended mind that, nevertheless, is at a certain position in space; or perhaps it is the brain which quite certainly is spatially located. It is still the case that, in entering into the relation of sensing with its sense-data, the latter are sensed absolutely objectively in the spatial relations they in fact have.

Amazingly, Jackson neither sees nor senses a difficulty here. This is shown by his criticism of the 'two-dimensional' view of sense-data which regards them as squashed flat in a plane at right-angles to the line of sensing.[1] It is worth quoting him at length here:

> A round plate held at a reasonably steep angle does *not* look the same as one held at right-angles to the line of vision. But one held at right-angles manifestly looks round. Therefore, the plate at an angle cannot be said to look round *simpliciter*. But it is also true that the plate held at an angle does *not* look the same as an elliptical plate held at right-angles to the line of vision. The latter, however, manifestly *does* look elliptical. Therefore the round plate at an angle cannot be said to look elliptical *simpliciter*. The dilemma facing the two-dimensional theorist is now clear: Is the sense-datum belonging to the round plate held at an angle, round or elliptical? If he opts for it being round, he fails to adequately reflect in his theory the differences between seeing the round plate straight on and seeing it at an angle; if he opts for it being elliptical, he fails to reflect the difference between seeing the round plate at any angle and seeing an elliptical plate straight on.

The two-dimensional theorist should, on logical grounds, reject this supposed dilemma. It follows from the relation the sensor has, in sensing, to his sense-data that the latter must have the properties they appear to have. If the sense-data from the round plate at an angle do not look round and if they do look some shape and the only other shape they can look is elliptical, then they are elliptical. The introduction of the new phrase 'looking round *simpliciter*' cannot alter the soundness of this argument—unless the idea is that the second and third premisses of the above argument are to be replaced with a notion of 'looking round *non-simpliciter*', resulting in the conclusion that the sense-data have the mysterious property of being round *non-simpliciter*.[2]

Jackson's own account of how his theory escapes the dilemma he thinks confronts the two-dimensional view is as follows:

> The sense-datum belonging to the round plate held at an angle is round *and at an angle*. Hence, it differs from the sense-datum belonging to an elliptical plate in being round, not elliptical (p. 104).

The spatial *relations* a round sense-datum bears to other things, including the sensor, can have no bearing on how its *qualities* are sensed. If there are in the same visual field two circular sense-data, touching at one point and at an angle of thirty degrees to each other, then in what relevant respect to they differ so that one sensed roundness is not quite the same as the other? One should sense two perfectly straightforward circles on either side of a thirty degree angle. It is no argument to say that no-one has ever sensed such a thing, for it follows from the basic principle of the theory of sensing (that sense-data are sensed just as they are), plus Jackson's view that circular sense-data in the one visual field can have such relations to each other, that the thing is perfectly possible and should be easy to arrange. All we need are a couple of circular plates, a decently lit room and somebody to do some looking.

---

[1] I will, from now on, use this phrase 'line of sensing', rather than 'line of sight', as conveying what Jackson really means.

[2] The premisses would have to be replaced, and not merely supplemented, otherwise, of course, the sense-datum would turn out to be both elliptical *and* round *non-simpliciter*.

This is to make the point, once more, that there can be no sensing of a sense-datum at an angle to a line of sensing even though this is what Jackson evidently means when he speaks of sense-data being at an angle. If we do retain the view that we not only *see* things from a point of view —i.e. at certain distances from, and angles to, us—but also similarly *sense* sense-data from a point of sensing, then we cannot stop the appearance–reality distinction once more breaking out. For example, consider what seems a bit like an elliptical sense-datum sloping away from one (as the sensor) but also slowly rotating until finally it is perpendicular to the line of sensing (perhaps the rotation is achieved by psychokinesis). As it rotates (and assuming it is subject to no deforming influences) the appearance of this sense-datum must also change just because the angle it presents to the sensing is changing. It will become rounder and rounder until, perhaps, it appears circular when sensed face-on.[3] Do not these new appearances then require, for their explanation, to be identified with new sense-data and so on to infinity?

A referee for ANALYSIS has objected that, just like rainbows, sense-data are "odd objects" and that it no more makes sense to speak of rotating the latter than of trying to walk around the former. However, the respect in which rainbows and sense-data are odd objects lies, according to the referee, in the supposed fact that they are not *objects* at all. Both rainbows and sense-data get "explained away" in terms of the *relations* between other things—rainbows by the relations between the sun, raindrops and the eyes and sense-data by those between "real things", light and eyes. Whatever the case for rainbows, I deny that this is true of sense-data. Indeed, I think Jackson would explain the strange properties of rainbows by invoking the idea of sense-data as objects. That he does regard them as objects having intrinsic properties could be shown by many references to his work; but I will content myself with the following quotations from p. 129 where he uses a Berkeleyan argument to go from 'Colour is subjective in that it qualifies mental items, while being an *intrinsic* quality of those items' (his italics) to the conclusion that 'sense-data have colour, *shape* and *extension*' (my italics). It must make sense to say of such objects that, if they have a certain orientation to a line of sensing, then they could have a different orientation; and this, as I argue, means that the appearance–reality distinction is once more involved.

---

[3] And if this is true then it is not clear that its colour may not also appear to change—may not the colour appear lighter when we sense the colour expanse face-on?

We should say, then, contrary to Jackson's account that the sense-datum belonging to a plate held at an angle *cannot* be circular (ignoring here complications arising from "perceptual constancy") and there is, for most angles, nothing else it can be but an ellipse. The point can be given added force if we consider rectilinear figures, rather than smoothly curved ones such as circles and ellipses, for I take it that Jackson also thinks that the sense-datum associated with, say, a rectangle held at an angle is rectangular. In particular, consider a long, thin rectangle sloping away from one so that somewhere in the distance two of its sides appear to meet—the effect one may get with railway lines extending away on a plain. The appearance is of a closed figure bounded by *three* straight lines so that, if the appearance is the reality, what we have is a closed, three-sided, rectilinear figure, i.e., we have a triangular sense-datum, not a rectangular one. Moreover, even though the sense-datum is undoubtedly triangular—it conforms impeccably to the definition of a triangle—it is not like seeing a triangle straight on. Therefore, there can be a triangular sense-datum belonging to a surface at an angle even though this sense-datum is not like the sense-datum belonging to a triangular surface facing one. Once we have convinced ourselves of the truth of this it will not, I think, take much to persuade us that, as the rectangle is gradually raised, the sense-data belonging to it form a series of trapezia grading from the nearly triangular to the nearly rectangular—even though none of these trapezia look like the sense-data from a trapezium seen straight on. And a similar moral can be drawn concerning the elliptical sense-data associated with round plates at angles.

If the sense-datum from a round plate at an angle is elliptical how do we account for the 'difference between seeing the round plate at any angle and seeing an elliptical plate straight on'? It must be said at once that the difference cannot be sought in the shapes, themselves, of the two sense-data. They are both elliptical, in just one sense of that word, and there is no reason why they should not be exactly the same shape. The only way to account for the difference is to seek it in some other property or properties—either intrinsic or relational—in which they differ. The proponent of the two-dimensional theory might seek it in, for example, the different colours, or shading, of the two identically elliptical sense-data; and he might supplement his account with, say, a theory of 'unconscious inference' whereby one infers from the colour, or whatever, of the sense-datum the position in three-dimensional space of the relevant physical surface. However, Jackson points out that it is a fact that sense-data are arranged in three-dimensional space. This means that not only does this allow an answer to the question above in terms of the different spatial relations the otherwise identical sense-data have, it compels it to take this form.

In opposition to Jackson's view, that 'The sense-datum belonging to the round plate held at an angle is round and at an angle', we should hold that the sense-datum (i) is *elliptical* and (ii) is at an angle. In affirming (ii), however, we are not saying, as Jackson wants to say, that the sense-datum is sensed from a point of sensing but that, for example, it is at an angle of thirty degrees to the round sense-datum. The round sense-datum is, then, of course, similarly at an angle of thirty degrees to the elliptical one and there is no question of the former being, as it were, sensed head-on. If the round sense-datum were to be rotated without distortion , then we would continue to sense it (if we do go on sensing it) as being a round sense-datum which takes up successively different angles to the elliptical sense-datum and may come to lie in the same plane as the latter. (That the last situation does not eventuate must then be taken as following from the brute fact, as pointed out below, that round sense-data exist only in certain planes in the visual field). The difference between otherwise identical sense-data lies, to put the case in its general form, in their different spatial positions and orientations and it is this which is sensed without distortion and accounts for the sensed difference.

If the above argument is sound then there is no question of refuting it by an appeal to phenomenology. However, if like Jackson we wish to defend a representative theory of perception, it does raise the question of how we explain the undoubted fact that we *see* things from a point of view. It is here that an appeal to phenomenology might be made. One might argue, for example, that, although the various sense-data are sensed not as from a position but in the spatial relations they actually have, a certain series of parallel planes in sensed space is privileged in that it is only within these planes that circular, square, etc. sense-data occur. It is these sense-data, occurring in these privileged planes, which are taken to be physical surfaces directly facing one. Once one commences this kind of investigation there is a large number of phenomenological factors one might take account of, e.g. the different sizes of otherwise similar sense-data in different parts of sensed space, the relations between the sense-data of the 'body-image' and other sense-data, and so on. It seems clear, however, that, as much as the two-dimensional theory, the three-dimensional theory needs still to explain—by 'unconscious inference' or whatever—how the leap of 'interpretation', from sense-fields to seeing, occurs.

Macquarie University © A. Olding 1980

A SPEECH ACT DILEMMA

By C. R. Carr

In *Words and Deeds* (Oxford 1978) David Holdcroft offers an account of illocutionary acts intended to be immune from criticisms he makes of previous accounts, one of those accounts being the theory proposed by Schiffer in Chapter IV of *Meaning* (Oxford 1972). In this discussion I will argue that on one plausible reading Holdcroft's account itself fails to meet a central challenge it raises for Schiffer, and then explain that failure by noting a general dilemma facing similar accounts of illocutionary acts.

Holdcroft distinguishes uttering something with a certain force from performing that illocutionary act, and the major work done by his analysis is done by the account of illocutionary force. He offers the following as a sufficient condition for uttering something with a certain force.[1]

(8) S uttered x with a certain illocutionary force if S uttered x openly intending
- (i) to do so;
- (ii) that x should be related iconically or conventionally to, or otherwise be appropriately correlated with, the proposition that p;
- (iii) that A should recognize S's intention (i) and (ii);
- (iv) that the fact that S uttered x openly intending (i), (ii), and (iii) should in the context of utterance, C, which S believes to obtain, provide A with reason to think that (S intends A to think that q(p)) because . . . (Holdcroft, p. 149)

The account is later amended for other intended effects so that A might be provided with a reason to think that S wants him to comply with some request or answer a question or provide a reason actually to comply.

Before sketching the difficulty for (8), a difficulty relating to condition (iv), a few comments on various of its features would be in order. 'Openly intending' is proposed to handle the difficulties for which Schiffer proposed the use of 'mutual knowledge*', but as Holdcroft is sceptical about the concept he reverts to a revision which Grice made to handle difficulties in his analysis of speaker meaning (H. P. Grice, 'Utterer's Meaning and Intentions', *Philosophical Review* 1969, p. 159). Grice's revision explicitly precludes an intention on the speaker's part which A is to think S does not have. To require the speaker openly to intend is to say that all of the relevant intentions are intended to be recognized by the audience. Despite its apparent positive character, Holdcroft, like Grice and unlike Schiffer, has provided a negative formulation. Since there were counterexamples to Grice's earlier analyses which took the form (no matter how many (n) intentions the definition mentioned) of a hidden $n+1$ intention, the force of Holdcroft's proposal is to say there is no such hidden intention.

---

[1] S is a speaker, A his audience, 'p' a proposition, 'q(p)' a proposition which contains 'p'; 'q' may be null.

The requirement in (ii) that x be appropriately correlated with P does much concealed work. The correlation may be a conversational implication where A is intended to calculate from x to P in C. To make this explicit would require (ii) to be much more complex.

Condition (iv) is where the problem lies. It is formulated to rule out what Holdcroft views, correctly, as counterexamples to the status of the intended effect in Schiffer's analysis. The counterexamples result from Schiffer's requirement that in performing an illocutionary act S must intend the audience to believe that P or intend the audience to form the intention to act (for reasons that satisfy a certain description). The problem is that S need have no such intentions in performing an illocutionary act.[2] That this is the difficulty (iv) is to avoid is clear when in defending it Holdcroft says, 'But if (8) is correct, and S merely has to intend that the fact that he has uttered x with certain intentions should, in the circumstances, provide A with a reason for thinking that S intends A to think that q(p), then S need not in fact intend A to think that q(p)' (Holdcroft p. 150).

Now Holdcroft makes it clear that (8) only provides a sufficient condition for uttering something with a certain force. However, (iv) was formulated to handle a failure in the necessity of previous attempts—the quote above indicates that (iv) is correcting a failure in the necessity of the intention mentioned in previous accounts. The condition was included to cover problems with the necessity of other accounts, and must, if successful, meet the test which gave rise to it. So, to see if (iv) solves the difficulty it was intended to solve we must see whether the intention it includes is necessary. For this reason we can ignore for now the claim that (8) is meant only to give us a sufficient condition. In examining the failure of (iv) to provide a necessary condition we will see why the account fails to provide a sufficient condition.

It is clear that what is basic to the success of (iv) is just what it is to provide a reason to think one is wanted to believe or comply or answer and so on. Not every sense of 'provide a reason' will work. It cannot be the case that S must intend to provide a reason which actually will incline A to comply or to think that S wants him to comply. An order S gives which he intends A not to carry out, and knows A will not carry out, need not be primarily intended to provide A with a reason which will in fact motivate him appropriately. Evidence for the ambiguity of 'provide a reason' can be seen in the coherence of an utterance by a reluctant commander who orders A to do something only because he was himself ordered to, and says to A, 'I do not intend my order to provide you with a reason to attack, but I must, none the less, order you to attack.' There must be some sense of providing a reason other than actually attempting to get A to respond or the formulation of (iv) is inadequate. The difficulty is determining in what sense S might be said to provide a reason for A where he may clearly not intend to give any motivating reason for A at all.

---

[2] For the difficulties with such an analysis see C. R. Carr, 'Speaker Meaning and Illocutionary Acts', *Philosophical Studies* 1978, pp. 281–291.

One cannot simply do away with (iv). The reason being provided by the speaker very often does serve to individuate illocutionary force. The difference between ordering and begging someone to shut the door is not the action the audience is to perform, shutting the door. It is reasonable to suppose that the difference is the reason for compliance, authority versus pity possibly, each provides. Further, if we may restrict for the sake of convenience the discussion from here to imperatival acts, it is likely that very often that reason actually is intended to cause (be part of the cause of) A's complying. If providing a reason is part of the analysis of performing an illocutionary act, then in performing an act, the speaker must invoke some particular relation between the intended effect and the reason for complying. The question is how 'providing' is to be understood if S can perform a given act without intending A to believe or comply, and therefore not intending to provide a reason which really will motivate A to comply.

Let us call providing a reason which actually inclines A or is intended actually to incline A giving an active reason. Rather than an active reason S has provided a reason in some more formal sense of 'provide a reason'. Since we are nowhere told just what it is to provide a reason, but only what kinds of counterexamples it is intended to avoid, we must seek some minimal interpretation that meets these conditions (Carr p. 289 suggests one possible way of relating them). Since reasons are reasons for believing or acting, a plausible candidate is that S has provided something which in more normal or ordinary circumstances than the special cases described above will get A to believe or comply (I use 'get A to believe or comply' as a shorthand for Holdcroft's more complex phrasing). It appears that more than this will miss the 'non-standard' acts, and less will fail to tie reasons to actions at all.

Let us start explicating this interpretation of 'providing a reason' as follows. Certain reasons can be invoked to get people to believe, act, and so on. Appealing to someone's pity is a way of getting him to do something, and pleading is a way of appealing to pity to get compliance. We normally expect that, when someone utters a sentence associated with some reasons-identifiable act, he is providing that reason for compliance (or belief in the case of non-imperative acts). So, if on a particular occasion S utters, say, 'I order you to leave!', it is reasonable to suppose that he had intended his authority as an active reason for the audience to have for leaving. Of course, the normal relation between reason, r, and intended effect need not exist, that is, need not be intended by S. In the case of a reluctantly given order S can set the stage so that it is certain to A that S is giving the order only because he has to, and still be ordering even though he does not intend A to comply. But as a way of providing a reason, the conventionalization of the speech act has a life of its own, and the normal relation between r and the intended effect will be brought to A's attention. A will be expected to see that in the absence of evidence to the contrary, evidence that there are other features which override S's intending to provide an active reason, A has been provided with an active reason. In giving the above order S is not providing an active reason for A. He is, though, giving the order because an identifying relation between S's authority and compliance could still be claimed to hold. Reasons still play a role in the analysis. A, in recognizing the act as an order, sees that a form of reasons has been brought to his attention which in normal circumstances would be intended to provide him with an active reason.

S's intention enters in this way. S knows A will see that S has brought to A's attention that in performing the act S has invoked the reasons normally invoked in performing such an act via the act through which just that form of reasoning can be invoked. If S intends the act to be recognized as an order, he will intend A to recognize that authority has been appealed to. What S intends, though, need not be providing an active reason. What S must intend, though, if the act is a reasons-identifiable one is that A will see that he has been provided with something which, in the absence of evidence to the contrary, he must take as intended to provide an active reason. So, even though the act is not intended to provide an active reason, it is still an act of reporting, pleading, ordering, questioning, and so on. This, then, seems a way of explicating 'providing a reason' which will square with the existence of the non-standard cases which need to be covered, but which will allow intending even in these cases to perform the requisite illocutionary act. There is no other obvious way of tying the reason to the belief or action which will meet this condition.

But if one can intend the relation between r and the intended effect not to be active, and still the act performed is the one identified by the form of r normally associated with that act, the account faces a serious problem. While it seems the notion of providing a reason cannot be stronger than this explanation of condition (iv) if it is to be necessary, there are other considerations which suggest it must be.

Suppose, for example, S makes a joke in the form of giving an order where he clearly intends not to be ordering. Conditions (i), (ii), and (iii) for the performance of the speech act S was not performing would be satisfied even in a joke. Suppose S jokingly says to a subordinate who has made some minor mistake, 'I order you to shoot yourself'. S intends (i) to utter something, 'I order you', which is (ii) conventionally or iconically or otherwise correlated with ordering, in this case conventionally. If it were not so correlated, the joke would not work. And (iii), S intends the audience to recognize (i) and (ii). But it also appears that (iv) would be satisfied, and if this is so, then S has ordered. It would be satisfied in the same way it is satisfied in those non-standard orders where S does not intend to provide an active reason. If S clearly intended this to be a joke, it is wrong to say that he ordered the subordinate to shoot himself, but for the joke to work at all S must invoke the relationship which normally holds between the reason in ordering and compliance. A must be intended to realize that in the absence of overriding considerations, here S's intention to be joking, S has done something which would provide A with an active reason. If S can order A by appealing to a standard relationship which is in that particular case not active, it seems that in the case of joking S likewise must be said to have ordered. For this account to allow non-standard speech acts to be speech acts it must allow jokes, and similar speech acts, to be speech acts they were not intended to be.

One might reply that in a case like joking one does not intend to be ordering while with non-standard orders one does. The trouble is that according to (8) one has ordered even in that case. Given the formulation of (8) the only way to deal with utterances which are not to be taken literally is to block one of the conditions specifying S's intentions. All, however, are satisfied. Until we are given an account of 'providing a reason' which can handle the non-standard illocutionary acts, those performed in the absence of intentions to realize the intended effect specified—and such cases are not that unusual—the account has not done what it was intended to do. More generally, the difficulty for this account of illocutionary force can be stated as follows. Any account stronger than this account of 'providing a reason', one which moves closer to providing an active reason, fails to handle the non-standard acts. An account as weak as this one makes non-literal utterances literal.

The difficulty we are up against here can be formulated as a more general dilemma for most current accounts of performing an illocutionary act. If one cashes out performing an illocutionary act in terms of a Gricean analysis of speaker meaning, one is faced with the problem that probably any illocutionary act can be performed in the absence of intentions to create psychological states such as intentions or beliefs or wants in the audience including beliefs about S's intentions. Non-problematic counterexamples abound. On the other hand various illocutionary acts do seem to be distinguished by the different reasons they are intended to provide the audience for intending to act or believing something or believing that the speaker believes something. The reasons help make them the specific acts they are. In the absence of any obvious way of mediating the conflicting tensions something, it would appear, must give.

Arkansas State University © C. R. CARR 1980

McHARRY'S THEODICY: A REPLY

By R. K. PERKINS, Jr.

A PROVOCATIVE piece by John D. McHarry (ANALYSIS 38.3, June 1978) attempts to provide a theodicy against the following argument from evil:

(1) God exists and is all good, all knowing, and all powerful.
(2) If God exists, the world must be in accord with His wishes, i.e., the best of all possible worlds.
(3) Therefore, this *is* the best of all possible worlds.
(4) This is *not* the best of all possible worlds; it contains too much evil.
(5) Therefore, (1) is false; there is no God.

McHarry's strategy—which eschews the traditional attempts to challenge line (4) via free-will, etc.—is to challenge the covert assumption that there is but one actual world among all those which are possible. He claims that there might be other actual worlds not spatiotemporally related to this one. Such a move allows him to hypothesize that God created the best possible world—or perhaps an infinite set of optimal worlds—and some other inferior, though positively valuable, worlds besides. The problem of evil is resolved by pointing out that our world may be one of those inferior, though positively valuable, worlds. The reason that God didn't make our world optimal too is because, being bound by a strong version of Leibniz's law of the identity of indiscernibles, He ran out of optimal variations, and had to settle for the 'next best'.

But there is a nest of difficulties connected with McHarry's theodicy. To begin with, his notion of 'world' is unduly narrow. I should have thought that it's more in keeping with philosophical tradition to construe 'world'—not as a 'collection of spatiotemporally interrelated objects' (p. 132)—but simply as everything that ever does exist, whether spatiotemporally interrelated or not. Let us call this wider sense of 'world' 'the universe'; and let us reserve the expression 'world' for those collections of spatiotemporally interrelated objects which comprise the universe. Let us restate the original argument:

(1′) God exists and is all good, all knowing, and all powerful.
(2′) If God exists, the universe must in be accord with His wishes, i.e., the best of all possible universes.
(3′) Therefore, this universe (of which this world is an element) *is* the best of all possible universes.
(4′) This universe is *not* the best of all possible universes; it contains too much evil.
(5′) Therefore, (1′) is false; there is no God.

Now, even if the theist need not be troubled by this world's imperfection, he must—or at least McHarry would—hold that the universe (the collection of all actual worlds) is the best possible. In other words, McHarry must challenge the truth of (4′). His move is (would be) to say that the universe *is* the best possible, even though it contains, as an element, a world which is imperfect; God could not have made this world a better world because any improved model already exists, and Leibniz's law forbids Him to have two worlds just alike.

But apart from the long standing controversies surrounding Leibniz's law, is it reasonable to suppose that God has used up *all* variations of better worlds so that He could not have made this one better? Let's look at some implications of this view.

What McHarry is apparently trying to undermine are arguments for (4) like the following: In our world there is much evil. Right now (let us suppose) there is a dog suffering a momentary twinge of toothache. Given God's omnipotence—which presumably includes His ability to alter causal sequences—He could have created this world just as it is except for the slight diminution of the dog's toothache; otherwise same antecedents, same consequents. Such a variation of our world would certainly seem to be a more valuable world since it apparently differs only with respect to the amount of evil it contains. Now if God could have done this, it would be reasonable to conclude not only that such a world, if actualized, would be better than our world, but that such a world would be an element of a collection—a universe—which would be better than the actual universe of which our world is an element. But this means that the *universe*, not just our world, would not be the best possible; (4′) would be true, and McHarry's theodicy would be wrecked. So it's necessary for him to hold that God couldn't have made our world along the lines of the diminished toothache suggested above; it's forbidden by Leibniz's law since, in some other world, God already did it.

If we attempt to work out some details of the above example, we shall see some of the difficulties that beset McHarry's proliferation of actual worlds. Suppose we contemplate our world being improved by diminishing the dog's toothache just slightly, (say) a just noticeable difference. God can't do it, says McHarry; it's already been done. What about a possible world just like ours except that the dog's pain is just noticeably less *and* (say) his shadow is one square billionth of a centimetre smaller? (I choose shadow size because it allows of an infinite number of variations which are, *prima facie*, value free.) And would not such a world, if actualized, be better than ours? Still no good; it's already been done. At this point we might try the following tack: if there is such a world as the one just described, then it, or some other very much like it, must be an actual world in which the dog's pain is just noticeably less and which is otherwise *more like our own than any other*. Let's suppose that such a world is one in which the dog's pain is just noticeably less and his shadow is only one square trillionth of a centimetre smaller. Then a world just like that one, except with the dog's shadow less than one square trillionth of a centimetre smaller than it is in our world, would be a non-actual world that God could have actualized in place of ours. So, the universe is not the best possible; and, again, it looks as if (4′) is, after all, true.

Now I take it that the kind of response that McHarry would want to make here would be one which takes issue with the way in which we 'found' such a non-actual world. He could simply deny that there is any actual world 'more like our own than any other'. The set of actual worlds which are like our world is a non-denumerably infinite set whose differences form a continuum like the set of points on a line. For no actual world is there one which is *most like* it any more than for a given point on a line there is a point which is *most near* it.

But if there is such an infinite set of actual worlds, it becomes difficult to see what God's motive could be for creating our world. On McHarry's view it would seem that the value of the universe is constituted by the total value of two sets—the set of optimal worlds and the set of less than optimal, but positively valuable worlds. But if the latter set is infinite, it is difficult to understand in what sense of 'better' the universe is better for including our world. The problem here is analogous, I think, to the problem of understanding in what sense of 'larger' the number of objects in an infinite set can be made larger by adding another object to it.

Assuming the coherence of the notion of actual infinite collections (*pace* Brouwer and his school), we could say that the universe without our world would be a proper subset of the universe with ours—the latter contains at least one positively valuable thing that the former would not contain, and in *that* sense is better. But still it is doubtful that we could meaningfully say that the total value of the latter is greater than that of the former. And, so, it's doubtful that God, *qua* all good, would have wished to create the universe of which our world is an element.

But even if we can attach an appropriate sense to saying that the universe is better with our world than without it such that God would wish to actualize our world, still, it does not appear plausible to say that the universe, with our world, is the *best* possible, unless there is some maximum or upper limit to the amount of positive value which a universe could contain. Our discussion of the toothache suggests that there could be an infinite number of other worlds different from ours but equal to it in value. And this suggests, I think, that there is no such upper limit. Thus the phrase 'the best possible universe' would seem to be nonsense, for the same reason that the phrase 'the greatest cardinal number' is nonsense. Or rather, to put the point more carefully, the phrase fails to refer to anything. But then, once again, we seem forced to conclude that (4′) is true, and that McHarry's theodicy fails.

One might say that this is not McHarry's problem: after all, the phrase originally occurred in an argument which he is trying to refute. Still, the phrase is one which many philosophers and theologians have claimed to extract analytically from the concept of God. And this suggests, I think, not that there is something wrong with the argument from evil, but rather that the traditional notion of God may be incoherent.

New England College

© R. K. Perkins, Jr. 1980

LIFESAVING

By T. M. REED

SUPPOSE that on Monday two persons, A and C, are stranded on one rock and that another person, B, is stranded on a second rock. A lifeboatman hears of the predicament, and knows that he has only time to go to one rock before both are submerged by the tide. Neither rock is more difficult to reach than the other, and the lifeboatman knows nothing about the identities of the stranded parties. Which rock should he go to?

According to the Anscombe view,[1] it is morally indifferent which rock he goes to. We should always save one life rather than no lives, but the rescuer wrongs no one, and therefore commits no wrong, if he saves one life rather than two.

By contrast, Jonathan Glover[2] advocates a maximizing policy: other things being equal, we ought always to prefer to save a larger to a smaller number of people. The Anscombe view, he claims, rests on one of two curious assumptions: it presupposes either that saving lives consecutively is more worthwhile than saving lives simultaneously, or that the rightness of saving a life has nothing to do with any value placed on the life which is saved.

I shall question this claim, by showing Glover's argument to entail either that the maximizing policy also involves one of these assumptions, or that the Anscombe view is mistaken merely because it differs from the maximizing policy.

Glover's argument is as follows. Suppose that A is on one rock and B on a second rock, and that all other circumstances are as described in the first paragraph. We might all agree that here it is morally indifferent which rock the lifeboatman goes to. Suppose now that on the next day, Tuesday, a third person, C, is stranded on one of the two rocks (the other rock being unoccupied). Again we may all agree that it is of some moral importance that the lifeboatman should go and rescue him. 'But,' Glover argues, 'according to the Anscombe view, if C had been stranded with either A or B on Monday, the prospect of saving him as well as one of the others need not have been considered of any moral importance at all' (p. 209). Thus, Glover thinks, the Anscombe view rests on the unstated belief that the time at which someone's life is in jeopardy is relevant to the importance of saving it, or else depends on the supposition that the rightness of saving a life has nothing to do with its desirable outcome.

---

[1] See G. E. M. Anscombe, 'Who is Wronged?', *The Oxford Review*, 1967. 'The Anscombe view' refers solely to the position criticized by Glover (see next footnote); I make no judgment as to whether Glover's interpretation of Anscombe is in all respects accurate.

[2] Jonathan Glover, *Causing Death and Saving Lives* (Harmondsworth, Middlesex: Penguin Books, 1977), pp. 207–209.

Unfortunately, Glover fails to notice that the maximizing policy has similar implications. Suppose that on Monday both A and B are stranded on the first rock, and the second rock is unoccupied. Obviously the maximizing policy requires that the lifeboatman rescue A and B. Suppose that on Tuesday C is stranded on the second rock, the first rock being unoccupied. Clearly the maximizing policy requires that the lifeboatman rescue C. But, it may now be argued, if on Monday C had been stranded on the second rock, the prospect of saving him need not (indeed could not) have been considered of any moral importance at all according to the maximizing policy. On Glover's pattern of argument, then, we are forced to conclude that the maximizing policy involves the unstated belief that the time at which a life is in jeopardy is morally relevant, or the supposition that the rightness of saving a life has nothing to do with any value placed on the life which is saved.

To this it may be responded that, if A and B had not been on the first rock, the maximizing policy would have required the lifeboatman to rescue C on Monday. But then we must allow a similar response on behalf of the Anscombe view: if neither A nor B had been stranded on Monday, then the lifeboatman would have been required to rescue C on that occasion. (If on Monday A had been stranded on the first rock with B safe ashore, then both the Anscombe view and the maximizing policy would require that A be saved if joined by C, and that either A or C be saved if the latter should occupy the second rock.)

Now it is clear that time preference cannot finally be the issue here. For on both the Anscombe view and the maximizing policy, lives which should be saved in certain circumstances need not be saved in other circumstances. Obviously such circumstances may vary with different times. Hence, if Glover's argument were cogent, one could as easily charge both views with space preference. Thus (e.g.) on Monday C might be located on a rock in some different place, with its own lifeboatman and no competing lives at stake; on the Anscombe view C should be rescued, though his rescue would not be required if he were stranded with A on the first rock or with B on the second in the original location. Likewise the maximizing policy would require C's rescue in the second location, but not in the original location if (e.g.) he should be stranded on the second rock while A and B occupy the first.

The issue then is whether the Anscombe view implies that the rightness of saving a life has nothing to do with any value placed on the life which is saved. Now since the Anscombe view requires the saving of life in the choice situations described, it can hardly imply indifference to *all* lives which are *saved*. That it implies indifference to some of them is however suggested by the following consideration:[3] if A is on one rock with the other unoccupied, C's joining A does not provide an additional reason for the lifeboatman's going to that rock, whereas it does provide an extra reason according to the maximizing policy. But this does not entail that on the Anscombe view the rightness of rescuing C has nothing to do with the value of his life, any more than it entails that C's joining A deprives the latter's life of all value; it shows only that on the Anscombe view the prospect of saving A (or C) is a fully sufficient reason for rescuing him, i.e., that the lifeboatman does as he ought by saving A and those who might join him in the situations described.[4] The criticism must therefore be: the Anscombe view attaches insufficient weight to certain lives, viz. numerous lives in choice situations which *could* be saved but which, eccentrically, it does not require us to save. But since both the Anscombe view and the maximizing policy entail some loss of life in choice situations, the criticism reduces to this: the Anscombe view does not require that saving a larger number of lives be preferred to saving a smaller number. This, however, is merely to point out that the Anscombe view represents a denial of the maximizing policy. It is not therefore to identify a curious assumption of the view; it is simply to give a partial statement of the view (whose other leading component is the requirement that we should save one life rather than no lives).

Of course, the Anscombe view may itself be curious. Whether one reaches this conclusion will perhaps depend on issues not addressed here: the extent to which doing wrong presupposes that persons are wronged, the tenability of a consistently applied maximizing policy, and so on. By itself, however, Glover's argument from lifesaving fails to discredit the position.

University of Utah © T. M. Reed 1980

---

[3] The author is indebted to the editor of ANALYSIS for calling his attention to this type of consideration.

[4] There is an evident analogy with the pacifist's attitude toward killing: that throwing a hand grenade at Jones would kill him is a sufficient reason on pacifist grounds for condemning the act; but the fact that Smith's joining Jones would not provide an additional reason does not entail that the wrongness of killing both must be unrelated to the value of one life or the other.

YOU'LL REGRET IT

By I. L. Humberstone

ONE is often advised not to take a particular course of action by being told that if one were to take that action, one would regret it afterwards—in the sense of wishing that one had done something else (maybe what is popularly called 'doing nothing', maybe not) instead. I will show that this consideration, by itself, is not a good reason for refraining from performing the action concerned, by showing that it is sometimes rational to do something one knows one will regret. 'Rational' here is being used in its ordinary everyday sense, according to which it would be inconsistent to hold that there was a good reason for not performing an action which it was all the same rational to perform. I shall not be commenting on what relation the example to be presented suggests there might be between this ordinary sense and the customary reconstruction thereof in contemporary decision theory, in terms of maximizing expected utility (or some variation on this). In addition, those troubled by the possibility of a 'narrow' notion of rationality which is such as to allow for the possibility of a conflict between the rational course of action and what a person (morally) ought to do, should note that the features which might be thought to give rise to such a conflict may be assumed absent from the example to be given.

The example concerns betting on horse races, a form of conduct I shall assume it is not always irrational to indulge in. I shall assume not only that it is sometimes rational to place a bet—to win—on a horse, but that it is sometimes (indeed typically) rational to do this when the amount bet is less than the total amount of the money at one's disposal at the time of the bet. If this is so, then we have an example of an action which it is rational for one to perform, but which one knows that one will regret having performed. For suppose that the horse wins: then one will regret not having placed more money on it. On the other hand, if the horse doesn't win, then one will regret having placed any money at all on it. Either way, one will regret having placed precisely that sum of money on the horse that one did in fact place. Furthermore, one can know all this before placing the bet—and if the bet was a rational one, coming to know this (say by being exposed to the argument just given) cannot make it less rational to go in for the bet. (Of course, in making any bet under normal conditions, one knows that one *might* regret it; what is important about the present case is that one knows with certainty that one definitely *will* regret it.)

It might be objected that it is never rational to place bets on horses, perhaps because the evidence available at the time of the bet is never a reliable enough indicator of the outcome, or of the probability of any particular outcome, for one to have a justified belief about the odds at which it would be appropriate to place the bet. But this is surely not a necessary feature of the example, and we can certainly imagine cases in which it is entirely lacking (because any situation in which betting would be rational may be transformed into such a case by introducing variable stakes with proportionate payoffs). Or, again, it might be objected that what one did (what one *really* did (!)) was to place a bet on a certain horse, and if that horse wins, one will not regret have done *that* (i.e., having placed some bet or other on that particular horse). A detailed diagnosis of the mistake behind this objection would require a discussion of the individuation of actions as well as of the intentionality of the context '. . . regrets that ——' —but why bother? An analogy brings out the point. Mightn't you regret pulling a trigger? But according to the spirit of the present objection, we can say that what you did was to move your finger—and why would anyone regret moving a finger?

Well, if the above example works, it might be a useful sort of case to think about in connection with practical reasoning. Not that other cases don't arise. One sort would be the kind of case in which a person has a chance to do something to help a cause he knows he is about to be brainwashed into thinking misguided: here, it might be held that it is rational for him to provide his help even though he will regret having done so. But this case does involve a kind of change of heart, which makes it not at all surprising that it also involves rational action one knows one will regret. Or again, I have occasionally heard it said in discussions of Newcomb's problem, that it is rational to go for the both-boxes solution (on the usual sorts of grounds) but that one knows, or at least has overwhelmingly good reasons for believing, that one will then lose out, and so wish that one had chosen differently. But this certainly won't do as a clear and straightforward example of a case where it is rational to do something one knows one will regret because it is precisely this sort of consideration that one-boxers appeal to in defence of their alternative solution to the problem. At least, the example I have offered is a bit easier to think about.

Monash University

© I. L. Humberstone 1980

The Philosophical Quarterly

Contents of Volume 30, *No.* 119, *April* 1980

Subjectivism and Seriousness	*E. P. Brandon*
Tasks	*John C. Hall*
States and Performances: Aristotle's Test	*Daniel W. Graham*
Comparative and Non-comparative Justice	*Phillip Montague*
DISCUSSION	
The Recovery of Illocutionary Force	*Keith Graham*

Edited in the University of St Andrews by Bernard Mayo with the assistance of Christopher Bryant, Anthony Ellis and Stephen Read
Published for the Scots Philosophical Club by the University of St Andrews

Annual Subscriptions:

Individuals £4.00 ($10.50)	Institutions £7.00 ($17.50)
Single issues £1.25 or $3.00	Institutions £2.00 or $5.00

for orders placed directly with the Journals Manager, Scottish Academic Press, Ltd., 33 Montgomery Street, EDINBURGH EH7 5JX, Scotland.

THEORIA

A Swedish journal of philosophy
Edited by
Bengt Hansson and Krister Segerberg

Volume 44, part 2

Peter Gärdenfors Fairness without interpersonal comparisons
Ingemund Gullvåg The logic of assertion
J. F. A. K. van Benthem and **W. J. Blok** Transitivity follows from Dummett's axiom

THEORIA, which invites contributions in all fields of analytic philosophy, usually publishes three issues a year. Subscriptions and requests for back numbers should be sent to the publishers, *Mssrs C. W. K. Gleerup, Box 1205, 221 05 Lund 1, Sweden.* Annual subscription is Skr. 45:– (approximately U.S. $ 10).

MIND

Vol LXXXIX No 355 July 1980

A quarterly review of philosophy
edited by D.W. Hamlyn with the assistance of S.D. Guttenplan

Articles:
Vulgar Consequentialism: J. J. KUPPERMAN
Aristotle's Ethical Holism: A. W. PRICE
The Diversity of Morals: F. E. SNARE
Do We Have to Know Why We Are Justified in Our Beliefs?: G. N. SCHLESINGER
Cohabitation, Stuff and Intermittent Existence: M. B. BURKE

Discussions
On what Matters in Survival: N. MEASOR
'Exists' is a Predicate: S. READ
The Ideal of Sincerity: Notes on a Footnote: N. J. H. DENT
Interpretations of Quantifiers: C. J. F. WILLIAMS
On Being Morally Authoritative: K. NIELSEN
Theories of Nature and the Nature of Theories: P. ROTH

Book Reviews
Books Received

Orders to: Journals Department, Basil Blackwell Publisher Ltd.,
108 Cowley Road, Oxford OX4 1JF, England
Annual subscription:
Institutions: £10.00 or $22.50 (Canada $26.00) *Individuals:* £8.00 or $18.00
(Canada $21.00) *Single issues:* £3.00 or $6.75 (Canada $7.75)

Published for the MIND ASSOCIATION by BASIL BLACKWELL PUBLISHER

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, J. H. Benson; T. R. Baldwin, Margaret A. Boden, A. E. Fisher, Andrew Harrison, R. F. Holland, Hidé Ishiguro, J. Kemp, J. H. McDowell, Bernard Mayo, D. H. Mellor, R. G. Swinburne, A. R. White, C. J. F. Williams, Peter Winch. This committee is responsible for appointing and advising the Editor and for the general policy of the paper.

SUBSCRIPTIONS. The subscription to ANALYSIS for institutions is £8.00 (inland), £9.80 (overseas), $24.00 (U.S.A.), $28.00 (Canada); for individuals £5.00 (inland), £6.00 (overseas), $15.00 (U.S.A.), $17.50 (Canada). Each volume comprises four numbers, three of 48 pages and one of 64 pages, appearing within the academic year—in January, March, June and October. Orders should be sent to Basil Blackwell, 108 Cowley Road, Oxford OX4 1JF, or placed with any bookseller.

CONTRIBUTIONS. **Articles submitted for publication should be addressed to Mr. Christopher Kirwan, Exeter College, Oxford OX1 3DP.** Contributors are asked to note the following requirements.

Articles should **normally** not exceed 3,000 words in length. **Occasionally,** however, longer contributions can be accepted;

They must be type written in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

It is regretted that owing to increased postal charges it is no longer possible to return typescripts unless the following instructions are followed:

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgement is required, a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope **and international reply coupons** of the requisite value, whether for air or surface mail.

Galley proofs of accepted articles will be sent to authors for correction, together with information about offprints. Typescripts will be retained by the Editor on the assumption that authors have kept their own copies.

The copyright of articles printed in ANALYSIS remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.

ISSN 0003–2638

27 MAY 1983

PRINTED IN GREAT BRITAIN BY BURGESS & SON (ABINGDON) LTD., ABINGDON, OXFORDSHIRE